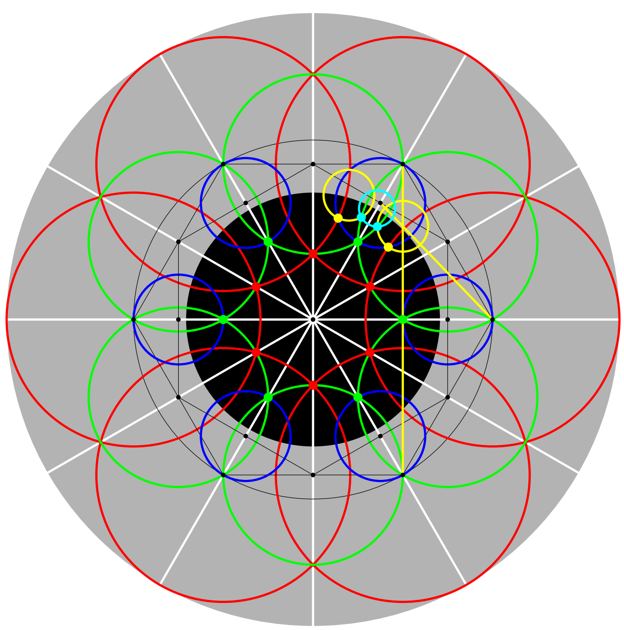
MATHEMATICS 322

Introduction to
ORDINARY DIFFERENTIAL EQUATIONS

Thomas Wieting

Reed College

2015



ORDINARY DIFFERENTIAL EQUATIONS

Chapter 0 OBJECTIVES AND PREREQUISITES

Objectives

1° Differential Equations state relations between functions and their derivatives. They provide a base for the analysis of Physical Systems. For illustrations, we point to the ubiquitous Potential Equation:

$$(\mathbf{S}) \qquad -\psi^{\circ\circ}(x) + V(x)\psi(x) = 0$$

which figures in Quantum Mechanics. We also point to the classical Equations of Newton governing Celestial Mechanics:

$$(\mathbf{N}) \qquad \begin{aligned} m_1\mathbf{r}_1^{\circ\circ}(t) &= \sum_{j\neq 1} \frac{m_j m_1}{\rho_{j1}(\mathbf{r}(t))^2}\frac{\mathbf{r}_j(t)-\mathbf{r}_1(t)}{\rho_{j1}(\mathbf{r}(t))} \\ m_2\mathbf{r}_2^{\circ\circ}(t) &= \sum_{j\neq 2} \frac{m_j m_2}{\rho_{j2}(\mathbf{r}(t))^2}\frac{\mathbf{r}_j(t)-\mathbf{r}_2(t)}{\rho_{j2}(\mathbf{r}(t))} \\ &\vdots \\ m_n\mathbf{r}_n^{\circ\circ}(t) &= \sum_{j\neq n} \frac{m_j m_n}{\rho_{jn}(\mathbf{r}(t))^2}\frac{\mathbf{r}_j(t)-\mathbf{r}_n(t)}{\rho_{jn}(\mathbf{r}(t))} \end{aligned} \qquad (0 < \rho(\mathbf{r}(t)))$$

and of Maxwell governing Electrodynamics:

$$(\mathbf{M}) \qquad \begin{aligned} (\nabla \bullet \mathbf{B})(t,x,y,z) &= 0 \\ \mathbf{B}_t(t,x,y,z) + (\nabla\times\mathbf{E})(t,x,y,z) &= \mathbf{0} \\ (\nabla \bullet \mathbf{E})(t,x,y,z) &= T(t,x,y,z) \\ -\mathbf{E}_t(t,x,y,z) + (\nabla\times\mathbf{B})(t,x,y,z) &= \mathbf{K}(t,x,y,z) \end{aligned}$$

One should note that (**S**) is a *simple* equation while (**N**) and (**M**) are *systems* of equations.

2° For equations (**S**) and (**N**), we have signaled differentiation with respect to the variables x and t by the superscript $\circ$. For the significance of the various symbols, see Chapters 2 and 4. For equation (**M**), we have signaled differentiation with respect to the variable t by the subscript t and we have signaled the action of the divergence and the curl in conventional manner. The variables $\mathbf{E}$, $\mathbf{B}$, T, and $\mathbf{K}$ denote the electric field, the magnetic field, the charge density, and the current density, respectively.

3° The functions which figure in a Differential Equation may depend upon one variable or upon many. In the former case, one refers to the equation as an

Ordinary Differential Equation (ODE), in the latter, as a Partial Differential Equation (PDE). Clearly, (**M**) is partial while (**S**) and (**N**) are ordinary.

4° It may happen that the family of functions which satisfy a Differential Equation form a linear space. We mean to say that, for any number c and for any solutions F, F', and F'' of the equation, cF and $F' + F''$ are also solutions. Clearly, (**S**) is linear. The homogeneous case of (**M**) (in which, by definition, $T = 0$ and $\mathbf{K} = \mathbf{0}$) is also linear.

Prerequisites

5° For smooth progress through the following Chapters, readers should be familiar with:

- (•) the basic methods of multivariable calculus
- (•) power series'
- (•) the concepts of linear space, matrix, and linear mapping

Notation

6° We employ the following notation:

$\mathcal{Z}$: the integers
$\mathcal{Q}$: the rationals
$\mathcal{R}$: the reals
$\mathcal{C}$: the complex numbers

Introduction

1° In this chapter, we present the basic form for Ordinary Differential Equations (ODEs) and we prove the Fundamental Theorem. As an example, we describe the Lotka/Volterra Equation from mathematical biology. By simple adaptations, we show that the basic form incorporates all the forms of ordinary differential equations which arise in practice. To illustrate the adaptations, we describe several common examples from mathematical physics.

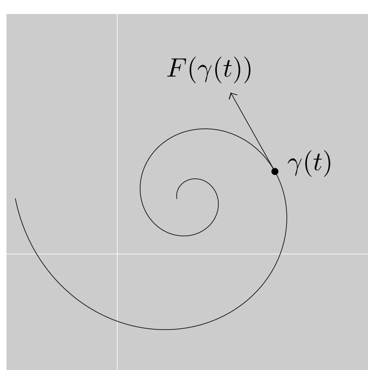
Autonomous First Order Ordinary Differential Equations

2° Let n be a positive integer. Let V be an open subset of $\mathcal{R}^n$ and let F be a mapping carrying V to $\mathcal{R}^n$. Let w be a member of V, let s be a number in $\mathcal{R}$, let J be an open interval in $\mathcal{R}$ containing s, and let γ be a differentiable mapping carrying J to $\mathcal{R}^n$ for which $\gamma(J) \subseteq V$. One says that γ is an *integral curve* for F passing through w at time s iff:

$$(\circ) \qquad \gamma^\circ(t) = F(\gamma(t)) \qquad (t \in J)$$

$$(\bullet) \qquad \gamma(s) = w$$

One refers to relation ($\circ$) as the Ordinary Differential Equation (ODE) defined by (the Velocity Field) F and to relation ($\bullet$) as an Initial Condition. One says that the ODE is Autonomous because F does not depend explicitly upon the time t and one says that it is First Order because nothing more than γ and γ° figure in it.



Integral Curve

3° One may express the relations (∘) and (•) in coordinates as follows:

$$(\circ) \qquad \begin{aligned} \gamma_1^\circ(t) &= F_1(\gamma_1(t), \gamma_2(t), \,\dots\,, \gamma_n(t)) \\ \gamma_2^\circ(t) &= F_2(\gamma_1(t), \gamma_2(t), \,\dots\,, \gamma_n(t)) \\ &\vdots \\ \gamma_n^\circ(t) &= F_n(\gamma_1(t), \gamma_2(t), \,\dots\,, \gamma_n(t)) \end{aligned} \qquad (t \in J)$$

$$(\bullet) \qquad \begin{aligned} \gamma_1(0) &= w_1 \\ \gamma_2(0) &= w_2 \\ &\vdots \\ \gamma_k(0) &= w_n \end{aligned}$$

Often, one adopts informal notation, such as the following:

$$(\circ) \qquad x^\circ = F(x) \qquad \text{or} \qquad \begin{aligned} x_1^\circ &= F_1(x_1, x_2, \,\dots\,, x_n) \\ x_2^\circ &= F_2(x_1, x_2, \,\dots\,, x_n) \\ &\vdots \\ x_n^\circ &= F_n(x_1, x_2, \,\dots\,, x_n) \end{aligned}$$

In such cases, one identifies the mapping γ with the vector variable x, which depends (implicitly) on t.

The Fundamental Theorem

4° The Fundamental Theorem for Autonomous First Order ODEs asserts that, for each number s in $\mathcal{R}$ and for each member w of V, there exists an integral curve $\hat{\gamma}$ for F passing through w at time s such that, for any integral curve γ for F passing through w at time s, γ is a restriction of $\hat{\gamma}$. That is, the domain J of γ is a subset of the domain $\hat{J}$ of $\hat{\gamma}$ and, for each t in J, $\gamma(t) = \hat{\gamma}(t)$. One refers to $\hat{\gamma}$ as the *maximum* integral curve for F passing through w at time s.

Proof of the Fundamental Theorem

5° We hasten to add that the Fundamental Theorem requires an hypothesis, which constrains the rate of change of F. Specifically, one requires that, for each member w of V, there are positive numbers r and c such that $B_r(w) \subseteq V$ and such that, for any members x and y of $B_r(w)$:

$$|F(x) - F(y)| \leq c\,|x - y|$$

(In this context, we take $B_r(w)$ to stand for the subset of $\mathcal{R}^n$ consisting of all members z such that $|z - x| \leq r$.) It would be necessary that F be continuous. It would be sufficient that F be continuously differentiable, but the more general requirement is useful.

6° Let us prove the theorem. For that purpose, we will apply the Contraction Mapping Theorem for complete metric spaces. Let s be a number in $\mathcal{R}$, let w be a member of V, let J be an open interval in $\mathcal{R}$, and let γ be a continuous mapping carrying J to $\mathcal{R}^n$ for which $\gamma(J) \subseteq V$. Obviously, γ is an integral curve for F passing through w at time s iff:

$$(*) \qquad \gamma(t) = w + \int_s^t F(\gamma(u))du \qquad (t \in J)$$

One should see in the foregoing relation a suggestion of a fixed point.

7° Let r, b, and c be positive numbers such that $B_r(w) \subseteq V$ and such that, for any members x, y, and z of $B_r(w)$:

$$|F(x)| \leq b \quad \text{and} \quad |F(y) - F(z)| \leq c\,|y - z|$$

Let σ be a positive number such that $\sigma b \leq r$ and $\sigma c < 1$. Let $\mathbf{X}$ be the family:

$$\mathbf{X} := \mathbf{M}\big((s - \sigma, s + \sigma), B_r(w)\big)$$

composed of all continuous mappings α carrying $(s - \sigma, s + \sigma)$ to $B_r(w)$. We may supply $\mathbf{X}$ with the uniform metric $\mathbf{m}$, as follows:

$$\mathbf{m}(\alpha_1, \alpha_2) := sup\{|\alpha_1(t) - \alpha_2(t)| \ : \ s - \sigma < t < s + \sigma\}$$

where α_1 and α_2 are any mappings in $\mathbf{X}$. By common knowledge, $\mathbf{X}$ is complete. For each α in $\mathbf{X}$, let β be the mapping carrying $(s - \sigma, s + \sigma)$ to $\mathcal{R}^n$, defined as follows:

$$\beta(t) := w + \int_s^t F(\alpha(u))du \qquad (s - \sigma < t < s + \sigma)$$

One can easily verify that β is in $\mathbf{X}$. Having done so, one may introduce the mapping $\mathbf{F}$ carrying $\mathbf{X}$ to itself, defined as follows:

$$\mathbf{F}(\alpha) := \beta \qquad (\alpha \in \mathbf{X})$$

One can easily verify that $\mathbf{F}$ is a contraction mapping. In fact, for any members α_1 and α_2 of $\mathbf{X}$, one can show that:

$$\mathbf{m}(\mathbf{F}(\alpha_1), \mathbf{F}(\alpha_2)) \leq \sigma c\,\mathbf{m}(\alpha_1, \alpha_2)$$

Consequently, by the Contraction Mapping Theorem, there is precisely one γ in $\mathbf{X}$ such that $\mathbf{F}(\gamma) = \gamma$. Obviously, γ is an integral curve for F passing through w at time s. The domain of γ is $(s - \sigma, s + \sigma)$.

8° By careful application of the foregoing result, one may proceed to prove the mature form of the Fundamental Theorem. Let us sketch the steps. First, one must prove that, for any number s in $\mathcal{R}$, for any open intervals J_1 and J_2 in $\mathcal{R}$, and for any integral curves γ_1 and γ_2 for F with domains J_1 and J_2, respectively, if $s \in J_1 \cap J_2$ and if $\gamma_1(s) = \gamma_2(s)$ then there is a positive real number τ such that $(s-\tau, s+\tau) \subseteq J_1 \cap J_2$ and such that the restrictions of γ_1 and γ_2 to $(s-\tau, s+\tau)$ coincide. Second, one must prove that, for any number s in $\mathcal{R}$, for any member w of V, for any open intervals J_1 and J_2 in $\mathcal{R}$, and for any integral curves γ_1 and γ_2 for F with domains J_1 and J_2, respectively, if $s \in J_1 \cap J_2$ and if $\gamma_1(s) = \gamma_2(s)$ then the restrictions of γ_1 and γ_2 to $J_1 \cap J_2$ coincide. Finally, one may prove the Fundamental Theorem. That is, one may prove that, for any number s in $\mathcal{R}$ and for any member w of V, there is a maximum integral curve γ for F passing through w at time s.

9° Just as well, we might set $q = (1/2)r$, we might select ρ so that $\rho b \leq q$ and $\rho c < 1$, and we might replace w by any member v of $B_q(w)$. The foregoing argument would remain valid. We would obtain a fixed point γ for $\mathbf{F}$. Of course, γ would be an integral curve for F passing through v at time s, with domain $(s - \rho, s + \rho)$. We may infer that:

$$D := (s - \rho, s + \rho) \times B_q(w) \subseteq \Delta$$

One refers to D as a Flow Box for F.

The Flow

10° Let w be a member of V. Let γ_w be the maximum integral curve for F passing through w at time 0 and let J_w be the domain of γ_w. One defines the *flow domain* Δ for F as follows:

$$\Delta := \{\, (t, w) \in \mathcal{R} \times V \ : \ w \in V, \ \ t \in J_w \,\}$$

By the preceding article, it is plain that Δ is an open subset of $\mathcal{R} \times V$. In turn, one defines the *flow mapping* γ for F, carrying Δ to V, as follows:

$$\gamma(t, w) := \gamma_w(t) \qquad ((t, w) \in \Delta)$$

11° For any real number t, one may introduce the (open) subset V_t of V consisting of all members w for which $(t, w) \in \Delta$ and one may define the mapping:

$$\gamma_t(w) := \gamma(t, w) \qquad (w \in V_t)$$

carrying V_t to V. The mappings γ_t and γ_w emphasize different aspects of the flow mapping γ, by fixing t while w varies and by fixing w while t varies.

Escape to the Boundary

12° Let x be a member of V and let γ_x be the maximal integral curve for F passing through x at time 0, with domain:

$$J_x = (a_x, b_x) \qquad (-\infty \leq a_x < 0 < b_x \leq \infty)$$

We say that γ_x *future escapes* to the boundary of V iff, for each compact subset M of V, there is some τ in J_x such that:

$$\gamma_x([\tau, b_x)) \cap M = \emptyset$$

Let us assume that $b_x < \infty$. We contend that γ_x future-escapes to the boundary of V.

13° Let us *suppose*, to the contrary, that there is a compact subset M of V such that, for each τ in J_x, $\gamma_x([\tau, b_x)) \cap M \neq \emptyset$. Hence, we may introduce an increasing sequence:

$$t_1 < t_2 < \ \cdots \ < t_j < \ \cdots \quad \uparrow \ \ b_x$$

in J_x, converging to b_x, such that, for each index j, $\gamma_x(t_j) \in M$. Since M is compact, we may apply the Bolzano/Weierstrass Theorem. In effect, we may take the sequence:

$$\gamma_x(t_1), \gamma_x(t_2), \ \ldots \ , \gamma_x(t_j), \ \ldots$$

in M to be convergent:

$$\gamma_x(t_j) \longrightarrow w, \qquad w \in M$$

With reference to article 9°, we may introduce positive numbers q and ρ such that:

$$(-\rho, \rho) \times B_q(w) \subseteq \Delta$$

Obviously, for each y in $B_q(w)$, $(-\rho, \rho) \subseteq J_y$. That is, the maximal integral curve γ_y for F passing through y at time 0 must be defined at least on the open interval $(-\rho, \rho)$.

14° Let j be an index such that:

$$b_x - t_j < \rho \quad \text{and} \quad \gamma_x(t_j) \in B_q(w)$$

Let $\tau = t_j$ and let $y = \gamma_x(\tau)$. Let δ be the mapping carrying $(a_x, \tau + \rho)$ to $\mathcal{R}^k$, defined as follows:

$$\delta(t) := \begin{cases} \gamma_x(t) & \text{if } a_x < t < b_x \\ \gamma_y(t-\tau) & \text{if } \tau - \rho < t < \tau + \rho \end{cases}$$

One can easily verify that δ is an integral curve for F passing through x at time 0. However, $b_x < \tau + \rho$, in contradiction with the definition of γ_x. We infer that our *supposition* is untenable. Therefore, if $b_x < \infty$ then γ_x future escapes to the boundary of V.

15° Of course, one may, in similar manner, formulate the concept of past escape to the boundary of V and one may prove that if $-\infty < a_x$ then γ_x past escapes to the boundary of V.

Convergence

16° We say that γ_x is *future convergent* iff there is a member y of V such that:

$$\lim_{t \to b_x} \gamma_x(t) = y$$

We refer to y as the future limit of γ_x. Let us assume that γ_x is future convergent. We contend that $b_x = \infty$ and that $F(y) = 0$.

17° To prove the first contention, we simply note that:

$$M := \gamma_x([0, b_x)) \cup \{y\}$$

is a compact subset of V. Consequently, γ_x does not future escape to the boundary of V. By the foregoing discussion, $b_x = \infty$.

18° In picturesque terms, one may say that if an integral curve future converges to a member of V then it must take infinitely long to do so.

19° To prove the second contention, we argue by contradiction. Let us *suppose* that $F(y) \neq 0$. Let $q = (1/2)|F(y)|$. Let r be a positive number such that $B_r(y) \subseteq V$ and such that, for each member z of $B_r(y)$, $F(z) \in B_q(F(y))$. Let τ be a number in J_x such that $\gamma_x([\tau, b_x)) \subseteq B_r(y)$. We find that, for each number t in (τ, b_x):

$$\frac{1}{t-\tau}(\gamma_x(t) - \gamma_x(\tau)) = \frac{1}{t-\tau}\int_\tau^t F(\gamma_x(u))du \in B_q(F(y))$$

Hence:

$$(1/2)(t-\tau)|F(y)| \leq |\gamma_x(t) - \gamma_x(\tau)|$$

It follows that $\gamma_x([\tau, b_x))$ is unbounded, in contradiction with our assumption that γ_x is future convergent. We infer that our *supposition* is untenable. Hence, $F(y) = 0$. Therefore, if γ_x is future convergent then $b_x = \infty$ and $F(y) = 0$, where y is the future limit of γ_x.

20° Of course, one may, in similar manner, formulate the concept of past convergence and one may prove that if γ_x is past convergent then $a_x = -\infty$ and $F(y) = 0$, where y is the past limit of γ_x.

21° One refers to a member y of V for which $F(y) = 0$ as a *critical point* for F.

Predator/Prey

22° Let a, b, c, and d be positive numbers. Let F be the mapping carrying $\mathcal{R}^+ \times \mathcal{R}^+$ to $\mathcal{R}^2$, defined as follows:

$$F(x_1, x_2) = (cx_1 - dx_1x_2, bx_1x_2 - ax_2) \qquad (0 < x_1, 0 < x_2))$$

The ODE defined by F is the ODE of Lotka and Volterra:

$$(\circ) \qquad \begin{aligned} x_1^\circ &= cx_1 - dx_1x_2 \\ x_2^\circ &= bx_1x_2 - ax_2 \end{aligned} \qquad (0 < x_1,\ 0 < x_2)$$

It serves to model the population dynamics of Prey (x_1) and Predator (x_2). Note that:

$$F(x_1, x_2) = (0, 0) \qquad \text{iff} \qquad x_1 = \frac{a}{b} \quad \text{and} \quad x_2 = \frac{c}{d}$$

Let h be the function defined as follows:

$$h(x_1, x_2) := bx_1 - a\,log(x_1) + dx_2 - c\,log(x_2) \qquad (0 < x_1,\ 0 < x_2)$$

One can easily verify that:

$$(\nabla h)(x_1, x_2) \bullet F(x_1, x_2) = 0 \qquad (0 < x_1, 0 < x_2)$$

Let γ be an integral curve for F:

$$\gamma(t) = (x_1(t), x_2(t)) \qquad (t \in J)$$

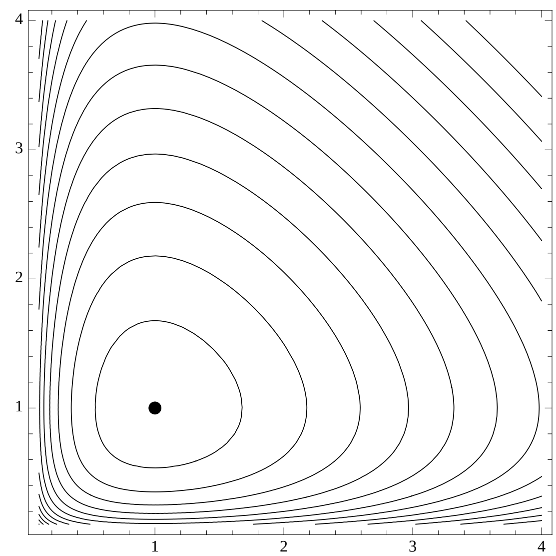
By the orthogonality relation just noted, it is plain that the function:

$$h(x_1(t), x_2(t)) \qquad (t \in J)$$

is constant. Consequently, $\gamma(J)$ is a subset of one of the level sets for h. For a sketch of the level sets for h, see the following figure. Obviously, the population pair:

$$(w_1, w_2) = (\frac{a}{b}, \frac{c}{d})$$

is critical. By interpreting the sketch, we find that, in general, the population pairs (x_1, x_2) evolve cyclically, in counterclockwise direction.



Phase Portrait for Predator/Prey

Approximate Solutions of ODEs

23° Let n be a positive integer. Let V be an open subset of $\mathcal{R}^n$ and let F be a mapping carrying V to $\mathcal{R}^n$. Let τ be a positive number. Let δ be a (piecewise continuously) differentiable mapping carrying the interval $[0, \tau]$ to $\mathcal{R}^n$ for which $\gamma(J) \subseteq V$. Let ϵ be a nonnegative number. We say that δ is an ϵ-*approximate* integral curve for F iff:

$$(\epsilon) \qquad \|\delta^\circ(t) - F(\delta(t))\| \leq \epsilon \qquad (0 \leq t \leq \tau)$$

24° For the (finitely many) values of t at which δ° admits a saltus, we mean to require that the foregoing inequality holds true for both the left and the right hand derivatives of δ.

25° Let us assume that there is a positive number c such that:

$$|F(x) - F(y)| \leq c\,|x-y| \qquad ((x,y) \in V \times V)$$

Under the assumption just stated, one can show that, for any ϵ'-approximate integral curve δ' for F and for any ϵ''-approximate integral curve δ'' for F:

$$\|\delta'(t) - \delta''(t)\| \leq \|\delta'(0) - \delta''(0)\| e^{ct} + \frac{\epsilon}{c}(e^{ct} - 1) \qquad (0 \leq t \leq \tau)$$

where $\epsilon = \epsilon' + \epsilon''$.

26° It may happen that δ'' is a true integral curve for F (so that $\epsilon'' = 0$) and that $\delta'(0) = \delta''(0)$. In such a case, we find that:

$$\|\delta'(t) - \delta''(t)\| \leq \frac{\epsilon'}{c}(e^{ct} - 1) \qquad (0 \leq t \leq \tau)$$

By the foregoing inequality, one may calculate the order of accuracy in the method of Runge and Kutta for designing approximate integral curves. We describe the method as follows.

The Method of Runge/Kutta

27° In the foregoing context, let us select a member w of V. Let n be a (large) positive integer and let $\sigma := \tau/n$. Let:

$$0 =: t_0 < \sigma =: t_1 < 2\sigma =: t_2 < \ldots < n\sigma =: t_n = \tau$$

be the partition of $[\,0, \tau\,]$ into n equal steps. One defines the *Runge-Kutta sequence* inductively as follows:

$$\begin{aligned} y_0 &:= w \\ z_1 &:= F(y_{j-1}) \\ z_2 &:= F(y_{j-1} + \frac{\sigma}{2} z_1) \\ z_3 &:= F(y_{j-1} + \frac{\sigma}{2} z_2) \\ z_4 &:= F(y_{j-1} + \sigma z_3) \\ y_j &:= y_{j-1} + \frac{\sigma}{6}(z_1 + 2z_2 + 2z_3 + z_4) \end{aligned} \qquad (1 \leq j \leq n)$$

Now one may define the polygonal curve δ as follows:

$$\delta(t) := y_{j-1} + \frac{1}{\sigma}(t - t_{j-1})(y_j - y_{j-1}) \qquad (1 \leq j \leq n,\ \ t_{j-1} \leq t \leq t_j)$$

This curve δ is the *Runge-Kutta approximation* to the integral curve γ for F passing through w at time 0. It turns out that:

$$sup\,\{|\delta(t)-\gamma(t)| : 0 \leq t \leq \tau\} \sim O(\sigma^4)$$

Reduction of Non-Autonomous First Order ODEs

28° Let n be a positive integer. Let V be an open subset of $\mathcal{R}^{1+n}$. Let $\mathcal{F}$ be a mapping carrying V to $\mathcal{R}^n$ and let F be the mapping carrying V to $\mathcal{R}^{1+n}$, defined in terms of $\mathcal{F}$ as follows:

$$F(u,x) := (1, \mathcal{F}(u,x)) \qquad ((u,x) \in V)$$

Let (s,w) be any member of V. Now we may consider the Non-Autonomous First Order Ordinary Differential Equation defined by $\mathcal{F}$:

$$x^\circ(t) = \mathcal{F}(t, x(t)) \tag{$\circ*$}$$

subject to the initial condition:

$$x(s) = w \tag{$\bullet*$}$$

and we may consider the Autonomous First Order Ordinary Differential Equation defined (as usual) by F:

$$\begin{aligned} u^\circ(t) &= 1 \\ x^\circ(t) &= \mathcal{F}(u(t), x(t)) \end{aligned} \tag{$\circ$}$$

subject to the following special form of the initial condition:

$$(u(s), x(s)) = (s, w) \tag{$\bullet$}$$

Obviously, x satisfies equation ($\circ*$) iff (u,x) satisfies equation ($\circ$), since the conditions $u^\circ(t) = 1$ and $u(s) = s$ force $u(t) = t$. Moreover, the initial conditions for x match. Hence, the integral curves for $\mathcal{F}$ may be obtained from certain of the integral curves for F.

29° Of course, we must justify application of the Fundamental Theorem to equation ($\circ$). To that end, we assume that, for each member (s,w) of V, there are positive numbers r and c such that $(s-r, s+r) \times B_r(w) \subseteq V$ and such that, for each number t in $(s-r, s+r)$ and for any members x and y in $B_r(w)$:

$$|\mathcal{F}(t,x) - \mathcal{F}(t,y)| \leq c|x-y|$$

This condition for $\mathcal{F}$ provides just what is needed. It implies that:

$$|F(t,x) - F(t,y)| \leq c|x-y|$$

which, in light of the special form of the initial condition, proves sufficient to justify application of the Fundamental Theorem.

Reduction of Autonomous Second Order ODEs

30° Let n be a positive integer. Let V be an open subset of $\mathcal{R}^{2n}$. Let $\mathcal{F}$ be a mapping carrying V to $\mathcal{R}^n$ and let F be the mapping carrying V to $\mathcal{R}^{2n}$, defined in terms of $\mathcal{F}$ as follows:

$$F(x,y) := (y, \mathcal{F}(x,y)) \qquad ((x,y) \in V)$$

Let s be any number in $\mathcal{R}$ and let (w, w°) be any member of V. Now we may consider the Autonomous Second Order ODE defined by $\mathcal{F}$:

$$(\circ\circ) \qquad x^{\circ\circ}(t) = \mathcal{F}(x(t), x^\circ(t))$$

subject to the initial condition:

$$(\bullet\bullet) \qquad (x(s), x^\circ(s)) = (w, w^\circ)$$

and we may consider the Autonomous First Order ODE defined (as usual) by F:

$$(\circ) \qquad \begin{aligned} x^\circ(t) &= y(t) \\ y^\circ(t) &= \mathcal{F}(x(t), y(t)) \end{aligned}$$

subject (as usual) to the initial condition:

$$(\bullet) \qquad (x(s), y(s)) = (w, w^\circ)$$

Obviously, x satisfies equation $(\circ\circ)$ iff (x, x°) satisfies equation $(\circ)$. Moreover, the initial conditions for x and x° match. Hence, the integral curves for $\mathcal{F}$ may be obtained from the integral curves for F.

31° To justify application of the Fundamental Theorem, we assume that, for any member (w, w°) of V, there are positive numbers r and c such that $B_r(w) \times B_r(w^\circ) \subseteq V$ and such that, for any members (x', y') and (x'', y'') of $B_r(w) \times B_r(w^\circ)$:

$$|\mathcal{F}(x', y') - \mathcal{F}(x'', y'')| \leq c\,|(x', y') - (x'', y'')|$$

One can easily check that the Fundamental Theorem would apply to equation $(\circ)$, hence to equation $(\circ\circ)$.

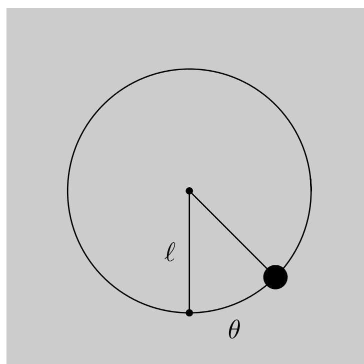
The Simple Pendulum

32° Let ℓ (the length of the pendulum) and g (the acceleration due to gravity at the surface of the Earth) be positive numbers. Let $\omega = \sqrt{g/\ell}$. One refers to ω as the natural frequency of the pendulum. Let $\mathcal{F}$ be the mapping carrying $\mathcal{R}^2$ to $\mathcal{R}$, defined as follows:

$$\mathcal{F}(\theta, v) = -\omega^2 sin(\theta) \qquad ((\theta, v) \in \mathcal{R}^2)$$

and let F be the corresponding mapping carrying $\mathcal{R}^2$ to $\mathcal{R}^2$:

$$F(\theta, v) = (v, -\omega^2 sin(\theta)) \qquad ((\theta, v) \in \mathcal{R}^2)$$



Simple Pendulum

Now equations (∘∘) and (∘) take the form:

(∘∘)
$$\theta^{\circ\circ}(t) = -\omega^2 sin(\theta(t))$$

(∘)
$$\begin{aligned} \theta^{\circ}(t) &= v(t) \\ v^{\circ}(t) &= -\omega^2 sin(\theta(t)) \end{aligned}$$

Let h (the energy per unit mass) be the function defined as follows:

$$h(\theta, v) = \frac{1}{2}v^2 - \omega^2 cos(\theta) \qquad ((\theta, v) \in \mathcal{R}^2)$$

One can easily verify that ∇h and F are orthogonal on $\mathcal{R}^2$:

$$(\nabla h)(\theta, v) \bullet F(\theta, v) = 0 \qquad (\theta, v) \in \mathcal{R}^2)$$

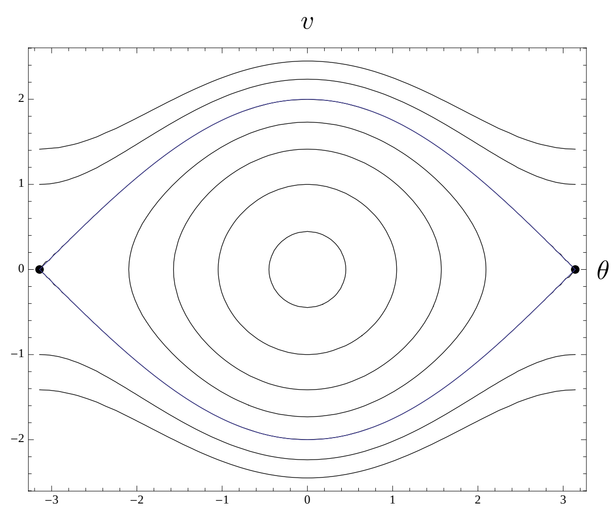
Let γ be an integral curve for F:

$$\gamma(t) = (\theta(t), \theta^{\circ}(t)) \qquad (t \in J)$$

By the orthogonality relation just noted, it is plain that the function:

$$h(\theta(t), \theta^{\circ}(t)) \qquad (t \in J)$$

is constant. Consequently, $\gamma(J)$ is a subset of one of the level sets for h. For a sketch of the level sets for h, see Figure 3. By interpreting the sketch, one can describe the relation between the initial conditions and the form of the corresponding integral curves.



Phase Portrait for Simple Pendulum

Newton

33° Let M (the mass of the Sun) and G (the gravitational constant) be positive numbers. Let $\mathcal{F}$ be the mapping carrying $V = (\mathcal{R}^3 \backslash \{\mathbf{0}\}) \times \mathcal{R}^3$ to $\mathcal{R}^3$, defined as follows:

$$\mathcal{F}(x, v) = -\frac{GM}{|x|^3} x \qquad ((x, v) \in V)$$

and let F be the corresponding mapping carrying V to $\mathcal{R}^3 \times \mathcal{R}^3$:

$$F(x, v) = (v, -\frac{GM}{|x|^3} x) \qquad ((x, v) \in V)$$

Now equations (∘∘) and (∘) take the form:

$$(\circ\circ) \qquad x^{\circ\circ}(t) = -\frac{GM}{|x(t)|^3}x(t)$$

$$(\circ) \qquad \begin{aligned} x^{\circ}(t) &= \quad v(t) \\ v^{\circ}(t) &= -\frac{GM}{|x(t)|^3}x(t) \end{aligned}$$

Equation (∘∘) is the simplest form of the Gravitational Equation of Newton.

34° Let h (the energy per unit mass) be the function defined as follows:

$$h(x,v) = \frac{1}{2}|v|^2 - \frac{GM}{|x|} \qquad ((x,v) \in V)$$

One can easily check that ∇h and F are orthogonal on V:

$$(\nabla h)(x,v) \bullet F(x,v) = 0 \qquad ((x,v) \in V)$$

Let γ be an integral curve for F:

$$\gamma(t) = (x(t), x^{\circ}(t)) \qquad (t \in J)$$

By the orthogonality relation just noted, it is plain that the function:

$$h(x(t), x^{\circ}(t)) = \frac{1}{2}|x^{\circ}(t)|^2 - \frac{GM}{|x(t)|} \qquad (t \in J)$$

is constant.

Lorentz

35° Let m (the mass of a particle) and q (the charge) be positive numbers. Let U be an open subset of $\mathcal{R}^3$ and let $V = U \times \mathcal{R}^3$ be the corresponding open subset of $\mathcal{R}^6 = \mathcal{R}^3 \times \mathcal{R}^3$. Let E (the electric field) and B (the magnetic field) be mappings carrying U to $\mathcal{R}^3$. Let $\mathcal{F}$ be the mapping carrying V to $\mathcal{R}^3$, defined as follows:

$$\mathcal{F}(x,y) = \frac{q}{m}(E(x) + y \times B(x)) \qquad ((x,y) \in V)$$

Now equations (∘∘) and (∘) take the form:

$$(\circ\circ) \qquad x^{\circ\circ}(t) = \frac{q}{m}(E(x(t)) + x^{\circ}(t) \times B(x(t)))$$

$$(\circ) \qquad \begin{aligned} x^\circ(t) &= y(t) \\ y^\circ(t) &= \frac{q}{m}(E(x(t)) + y(t) \times B(x(t))) \end{aligned}$$

One refers to equation $(\circ\circ)$ as the Lorentz Force Equation. It defines the motion of a charged particle in an electromagnetic field.

Reduction of Non-Autonomous Second Order ODEs

36° Let n be a positive integer. Let V be an open subset of $\mathcal{R}^{1+2n}$. Let $\mathcal{F}$ be a mapping carrying V to $\mathcal{R}^n$ and let F be the mapping carrying V to $\mathcal{R}^{1+2n}$, defined in terms of $\mathcal{F}$ as follows:

$$F(u,x,y) := (1, y, \mathcal{F}(u,x,y)) \qquad ((u,x,y) \in V)$$

Let (s, w, w°) be any member of V. Let us consider the Non-Autonomous Second Order Ordinary Differential Equation defined by $\mathcal{F}$:

$$(\circ\circ*) \qquad x^{\circ\circ}(t) = \mathcal{F}(t, x(t), x^\circ(t))$$

subject to the initial condition:

$$(\bullet\bullet*) \qquad (x(s), x^\circ(s) = (w, w^\circ)$$

and the Autonomous First Order Ordinary Differential Equation defined (as usual) by F:

$$(\circ) \qquad \begin{aligned} u^\circ(t) &= 1 \\ x^\circ(t) &= y(t) \\ y^\circ(t) &= f(u(t), x(t), y(t)) \end{aligned}$$

subject to the (specialized) initial condition:

$$(\bullet) \qquad (u(s), x(s), y(s)) = (s, w, w^\circ)$$

Obviously, x satisfies equation $(\circ\circ*)$ iff (u, x, x°) satisfies equation $(\circ)$, since the conditions $u^\circ(t) = 1$ and $u(s) = s$ force $u(t) = t$. Of course, the initial conditions for x and x° match. Hence, the integral curves for $\mathcal{F}$ may be obtained from certain of the integral curves for F.

37° To justify application of the Fundamental Theorem, one need only review the foregoing cases.

The Damped Forced Pendulum

38° Let q (the damping parameter), ω (the natural frequency), f (the forcing amplitude), and ν (the drive frequency) be positive constants. Let $\mathcal{F}$ be the mapping carrying $\mathcal{R}^3$ to $\mathcal{R}$, defined as follows:

$$\mathcal{F}(u,\theta,v) := -qv - \omega^2 sin(\theta) + f cos(\nu t) \qquad ((u,\theta,v) \in \mathcal{R}^3)$$

and let F be the corresponding mapping carrying $\mathcal{R}^3$ to $\mathcal{R}^3$:

$$F(u,\theta,v) := (1, v, -qv - \omega^2 sin(\theta) + f cos(\nu t)) \qquad ((u,\theta,v) \in \mathcal{R}^3)$$

Now equations $(\circ\circ *)$ and $(\circ)$ take the form:

$$(\circ\circ *) \qquad \theta^{\circ\circ}(t) = -q\theta^{\circ}(t) - \omega^2 sin(\theta(t)) + f\, cos(\nu t)$$

$$(\circ) \qquad \begin{aligned} u^{\circ}(t) &= 1 \\ \theta^{\circ}(t) &= v(t) \\ v^{\circ}(t) &= -qv(t) - \omega^2 sin(\theta(t)) + f\, cos(\nu u(t)) \end{aligned}$$

The integral curves for the forced damped pendulum provide a rich setting for the study of Chaotic Dynamics.

Chapter 2 SECOND ORDER LINEAR THEORY

Introduction

1° Let us consider the special case of First Order Linear ODEs. (For convenience of expression, we have dropped the adjective *NonAutonomous*.) We present just the basic definitions and facts. Then, by a simple adaptation, we concentrate upon the primary case of interest: the Second Order Linear ODEs in one variable ($n = 1$).

First Order Linear ODEs

2° Let n be a positive integer. Let I be an open interval in $\mathcal{R}$ and let $V = I \times \mathcal{R}^n$ be the corresponding open subset of $\mathcal{R}^{1+n}$. Let $\mathcal{F}$ be a continuous mapping carrying V to $\mathcal{R}^n$ such that, for each number t in I, $\mathcal{F}$ is linear in the variable member x of $\mathcal{R}^n$. In effect, we may present $\mathcal{F}$ as a matrix with n rows and n columns and with entries m_{jk} which are continuous functions of t:

$$\mathcal{F}(t, \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}) = \mathcal{F}(t) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} m_{11}(t) & m_{12}(t) & \cdots & m_{1n}(t) \\ m_{21}(t) & m_{22}(t) & \cdots & m_{2n}(t) \\ \vdots & \vdots & \vdots & \vdots \\ m_{n1}(t) & m_{n2}(t) & \cdots & m_{nn}(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

The ODE defined by $\mathcal{F}$ stands as follows:

$$(\circ *) \qquad \begin{aligned} \gamma_1^\circ(t) &= m_{11}(t)\gamma_1(t) + m_{12}(t)\gamma_2(t) + \cdots + m_{1n}(t)\gamma_n(t) \\ \gamma_2^\circ(t) &= m_{21}(t)\gamma_1(t) + m_{22}(t)\gamma_2(t) + \cdots + m_{2n}(t)\gamma_n(t) \\ &\vdots \\ \gamma_k^\circ(t) &= m_{n1}(t)\gamma_1(t) + m_{n2}(t)\gamma_2(t) + \cdots + m_{nn}(t)\gamma_n(t) \end{aligned} \qquad (t \in I)$$

The initial condition stands as usual:

$$(\bullet *) \qquad \begin{aligned} \gamma_1(s) &= w_1 \\ \gamma_2(s) &= w_2 \\ &\vdots \\ \gamma_n(s) &= w_n \end{aligned}$$

where s is a number in I and where w is a member of $\mathcal{R}^n$.

3° We may abbreviate the foregoing relations as follows:

$$(\circ *) \qquad \gamma^\circ(t) = \mathcal{F}(t)\gamma(t) \qquad (t \in I)$$

$$(\bullet *) \qquad \gamma(s) = w$$

4° Let us pause to verify that the Fundamental Theorem applies to $\mathcal{F}$. To that end, let u and v be numbers in $\mathcal{R}$ such that $[u, v] \subseteq I$. Let c be a number in $\mathcal{R}$ such that, for any indices j and k $(1 \leq j, k \leq n)$ and for any number t in $[u, v]$:

$$|m_{jk}(t)| \leq c$$

One can easily verify that, for any number t in $[u, v]$ and for any members x and y of $\mathcal{R}^n$:

$$|\mathcal{F}(t)(x - y)| \leq c|x - y|$$

Now article 29° in Chapter 1 justifies application of the Fundamental Theorem.

5° The foregoing observation yields a bonus. It shows that the domains of the various maximum integral curves for $\mathcal{F}$ must equal I. In fact, a counter instance γ would fail, in past or in future, to escape to the boundary of $I \times \mathcal{R}^n$, in contradiction with our discussion of these matters in articles 12°, 13°, 14°, and 15° in Chapter 1.

6° Let us denote by $\mathbf{G}$ the family of all maximum integral curves for $\mathcal{F}$. We will refer to such curves as *solutions* of $(\circ *)$. Obviously, $\mathbf{G}$ is a linear space over $\mathcal{R}$. We mean to say that, for any solutions γ, γ', and γ'' in $\mathbf{G}$ and for any number c in $\mathcal{R}$, $c\gamma$ and $\gamma' + \gamma''$ are also solutions in $\mathbf{G}$.

7° Let Λ be the mapping carrying $\mathbf{G}$ to $\mathcal{R}^n$, defined as follows:

$$\Lambda(\gamma) = \gamma(s) \qquad (\gamma \in \mathbf{G})$$

Obviously, Λ is linear. By the Fundamental Theorem, it is bijective. We infer that Λ is a linear isomorphism. Hence, the linear space $\mathbf{G}$, consisting of all solutions γ of $(\circ *)$, is n-dimensional.

8° The primary objective for a study of the Linear Theory is to describe a useful basis:

$$\beta_1, \beta_2, \ \ldots \ , \beta_n$$

of solutions in $\mathbf{G}$.

9° Let us turn to our objective. Let I be an open interval in $\mathcal{R}$ and let $V = I \times \mathcal{R}^2$ be the corresponding open subset of $\mathcal{R}^{1+2}$. Let p_0 and p_1 be continuous functions defined on I. Let $\mathcal{F}$ be the mapping carrying V to $\mathcal{R}^2$, defined as follows:

$$\mathcal{F}(t, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = \begin{pmatrix} 0 & 1 \\ -p_0(t) & -p_1(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad ((t,x) \in V)$$

Obviously, $\mathcal{F}$ is a special case of the general form introduced in article 2°, where:

$$\mathcal{F}(t) = \begin{pmatrix} 0 & 1 \\ -p_0(t) & -p_1(t) \end{pmatrix}$$

The ODE defined by $\mathcal{F}$ stands as follows:

$$(\circ *) \qquad \begin{aligned} \gamma_1^\circ(t) &= \quad \gamma_2(t) \\ \gamma_2^\circ(t) &= -p_0(t)\gamma_1(t) - p_1(t)\gamma_2(t) \end{aligned} \qquad (t \in I)$$

and the initial condition takes the form:

$$(\bullet *) \qquad \begin{aligned} \gamma_1(0) &= w \\ \gamma_2(0) &= w^\circ \end{aligned}$$

where s is a number in I and where w and w° are members of $\mathcal{R}$.

10° Let us replace γ_1 by f and let us restate equations $(\circ *)$ and $(\bullet *)$ as follows:

$$(\circ\circ *) \qquad f^{\circ\circ}(t) + p_1(t) f^\circ(t) + p_0(t) f(t) = 0 \qquad (t \in I)$$

$$(\bullet\bullet *) \qquad f(s) = w, \;\; f^\circ(s) = w^\circ$$

In this way, we obtain the conventional form for Second Order Linear ODEs in one variable.

11° For flexibility, we now encourage the values of our functions to be complex. By doing so, we lose nothing and we gain much. We continue to take the values of t to be real but, hereafter, we take the values:

$$p_0(t), \; p_1(t), \; f(t)$$

of p_0, p_1, and f to be complex and, consequently, we take the values of the initial conditions w and w° to be complex as well.

12° Now the family **G** of all solutions of $(\circ\circ *)$ is a linear space over $\mathcal{C}$. The dimension of **G** is two.

The Wronskian

13° Let h_1 and h_2 be solutions in **G**. We mean to describe a condition for determining whether or not:

$$h_1, h_2$$

is a basis for **G**. To that end, we form the *wronskian* of h_1 and h_2:

$$w(t) = det \begin{pmatrix} h_1(t) & h_2(t) \\ h_1^\circ(t) & h_2^\circ(t) \end{pmatrix} \qquad (t \in I)$$

One can easily verify that:

$$w^\circ(t) + p_1(t)w(t) = 0 \qquad (t \in I)$$

It follows that, for each s in I:

$$w(t) = exp(-\hat{p}_1(t))w(s) \qquad (t \in I)$$

where $\hat{p}_1$ is the antiderivative for p_1 such that $\hat{p}_1(s) = 0$.

14° Now it is plain that one or the other of the following two conditions must hold:

$(*)$ for each t in I, $w(t) = 0$
$(*)$ for each t in I, $w(t) \neq 0$

We contend that:

$$h_1, h_2$$

is a basis for **G** iff the second of the foregoing conditions holds.

15° To prove the contention, let us introduce a number s in I. We note that h_1 and h_2 are linearly dependent iff there exist numbers c_1 and c_2 in $\mathcal{C}$ such that:

$$|c_1|^2 + |c_2|^2 \neq 0 \quad \text{and} \quad \begin{pmatrix} h_1(s) & h_2(s) \\ h_1^\circ(s) & h_2^\circ(s) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

because, by the Fundamental Theorem, the condition just stated means that $c_1h_1 + c_2h_2$ is identically zero. Now one may complete the proof by routine observations.

Homogeneous/NonHomogeneous

16° Very often, one encounters the following generalization of the Second Order Linear ODE in one variable:

$$(\circ\circ *) \qquad f^{\circ\circ}(t) + p_1(t)f^{\circ}(t) + p_0(t)f(t) = \phi(t) \qquad (t \in I)$$

The function ϕ, presumed continuous, provides the new feature. To distinguish the old and new forms, one refers to the old form as the Homogeneous Case and to the new form as the NonHomogeneous Case.

17° The Damped Forced Pendulum provides an example. See article 38° in Chaper 1.

18° Remarkably, the new is contained in the old. We mean to say that, from a basis:

$$h_1, h_2$$

of solutions for the Homogeneous Case, one may, by quadrature, construct a particular solution g of the NonHomogeneous Case. That done, we may present the general solution of the NonHomogeneous Case as follows:

$$f = c_1 h_1 + c_2 h_2 + g$$

where c_1 and c_2 run through $\mathcal{C}$. This pattern is universal in the study of Linear Differential Equations.

19° One constructs g from h_1 and h_2 by *variation of constants*. That is, one forms a fake linear combination of h_1 and h_2:

$$g = u_1 h_1 + u_2 h_2$$

where the coefficients u_1 and u_2 are not constants but functions, at the moment unknown but soon to be determined. Requiring:

$$(1) \qquad u_1^{\circ} h_1 + u_2^{\circ} h_2 = 0$$

one finds that:

$$g^{\circ} = u_1 h_1^{\circ} + u_2 h_2^{\circ}$$

Requiring:

$$(2) \qquad u_1^{\circ} h_1^{\circ} + u_2^{\circ} h_2^{\circ} = \phi$$

as well, one finds that:

$$g^{\circ\circ} + p_1 g^{\circ} + p_0 g = \phi$$

To justify the procedure, one must justify the requirements (1) and (2). To that end, one rewrites the requirements as follows:

$$\begin{pmatrix} h_1 & h_2 \\ h_1^\circ & h_2^\circ \end{pmatrix} \begin{pmatrix} u_1^\circ \\ u_2^\circ \end{pmatrix} = \begin{pmatrix} 0 \\ \phi \end{pmatrix}$$

Since the wronskian w for h_1 and h_2 never vanishes, one may rewrite the requirements once again:

$$\begin{pmatrix} u_1^\circ \\ u_2^\circ \end{pmatrix} = \frac{1}{w} \begin{pmatrix} h_2^\circ & -h_2 \\ -h_1^\circ & h_1 \end{pmatrix} \begin{pmatrix} 0 \\ \phi \end{pmatrix} = \frac{\phi}{w} \begin{pmatrix} -h_2 \\ h_1 \end{pmatrix}$$

That is:

$$u_1^\circ = -\frac{\phi}{w} h_2, \quad u_2^\circ = \frac{\phi}{w} h_1$$

Now one justifies the procedure by quadrature.

One Solution is Sufficient

20° Let h_1 be a solution in $\mathbf{G}$ such that, for each t in I, $h_1(t) \neq 0$. We plan to design, by quadrature, another solution h_2 such that:

$$h_1, h_2$$

is a basis for $\mathbf{G}$. To that end, we form a multiple of h_1:

$$h_2 = u h_1$$

where u is a function specified, in steps, as follows:

(3) $$q \equiv 2\frac{h_1^\circ}{h_1} + 1, \quad v^\circ + q\, v = 0, \quad u^\circ \equiv v$$

We find that h_2 is a solution in $\mathbf{G}$. The wronskian for h_1 and h_2 stands as follows:

$$w = det \begin{pmatrix} h_1 & u h_1 \\ h_1^\circ & u^\circ h_1 + u h_1^\circ \end{pmatrix} = h_1(u^\circ h_1 - u h_1^\circ)$$

Given any number s in I, we may set the values $v(s) = u^\circ(s)$ and $u(s)$ so that:

$$u^\circ(s) h_1(s) - u(s) h_1^\circ(s) \neq 0$$

In this way, we insure that h_1 and h_2 are linearly independent.

Simplification of Form

21° Let u be a function defined on the interval I such that, for each t in I, $u(t) \neq 0$. Let f and g be functions defined on I for which:

$$f = ug \tag{4}$$

By routine computation, one can check that f satisfies the equation:

$$f^{\circ\circ} + p_1 f^{\circ} + p_0 f = 0 \tag{5}$$

iff g satisfies the equation:

$$g^{\circ\circ} + q_1 g^{\circ} + q_0\, g = 0 \tag{6}$$

where:

$$q_0 = \frac{1}{u}(u^{\circ\circ} + p_1 u^{\circ} + p_0 u) \;\text{ and }\; q_1 = \frac{1}{u}(2u^{\circ} + p_1 u)$$

Obviously, we may eliminate q_1 from equation (6) by choosing u so that:

$$u^{\circ} + \frac{1}{2} p_1 u = 0$$

By this observation and by the foregoing articles of discussion, we may present the equation:

$$-\psi^{\circ\circ}(t) + V(t)\psi(t) = 0 \tag{$\star$}$$

as an encapsulation of the theory of Second Order Linear ODEs in one variable.

22° With due respect to Mathematical Physics, we have inserted a significant minus sign in equation ($\star$), we have replaced g by the ubiquitous symbol ψ, and we have replaced q_0 by the conventional symbol V for what in practice proves to be a *potential function*.

Sturm Separation

23° Let h_1 and h_2 be linearly independent solutions of equation ($\star$). Let there be numbers s' and s'' in I such that $s' < s''$ and such that $h_2(s') = 0$ and $h_2(s'') = 0$. We contend that there must be some number t in I such that $s' < t < s''$ and such that $h_1(t) = 0$.

24° Of course, we may assume that the zeros s' and s'' of h_2 are successive, so that, for each number t in I, if $s' < t < s''$ then $h_2(t) \neq 0$. Hence:

$$h_2^\circ(s')h_2^\circ(s'') < 0$$

Since the wronskian w for h_1 and h_2 never vanishes, we find that $h_1(s')h_2^\circ(s')$ and $h_1(s'')h_2^\circ(s'')$ must have the same sign. Hence:

$$h_1(s')h_1(s'') < 0$$

Consequently, there must be some number t in I such that $s' < t < s''$ and such that $h_1(t) = 0$.

Sturm Comparison

25° Now let us consider two instances of equation $(\star)$:

$$-\psi_1^{\circ\circ}(t) + V_1(t)\psi_1(t) = 0$$
$$-\psi_2^{\circ\circ}(t) + V_2(t)\psi_2(t) = 0$$

Let s' and s'' be numbers in I for which $s' < s''$. Let us assume that, for each number t in I, if $s' < t < s''$ then:

$$V_1(t) \leq V_2(t)$$

and let us assume that there is some number s in I such that $s' < s < s''$ and:

$$V_1(s) < V_2(s)$$

Let h_1 be a nontrivial solution of the first of the foregoing equations and let h_2 be a nontrivial solution of the second. Let us assume that $h_2(s') = 0$ and $h_2(s'') = 0$. We contend that there must be some number t in I such that $s' < t < s''$ and such that $h_1(t) = 0$.

26° Again, we may assume that the zeros s' and s'' of h_2 are successive, so that, for each number t in I, if $s' < t < s''$ then $h_2(t) \neq 0$, indeed, $0 < h_2(t)$, so that $0 < h_2^\circ(s')$ and $h_2^\circ(s'') < 0$. In turn, let us suppose that, for each number t in I, if $s' < t < s''$ then $h_1(t) \neq 0$, indeed, $0 < h_1(t)$, so that $0 \leq h_1(s')$ and $0 \leq h_1(s'')$. Finally, ignoring the fact that h_1 and h_2 are solutions of distinct equations, let us form the wronskian w and compute its derivative:

$$w = h_1h_2^\circ - h_2h_1^\circ, \quad w^\circ = (V_2 - V_1)h_1h_2$$

27° Now we find that:

- $(*)$ $0 \leq w(s')$ and $w(s'') \leq 0$
- $(*)$ for each number t in I, if $s' < t < s''$ then $0 \leq w^\circ(t)$
- $(*)$ there is some number t in I such that $s' < t < s''$ and $0 \leq w^\circ(t)$

These conclusions are mutually contradictory. It follows that our prior supposition is false. Hence, there must be some number t in I such that $s' < t < s''$ and such that $h_1(t) = 0$.

Power Series Solutions

28° At this point, we replace the interval I by an open disk Δ in the complex plane $\mathcal{C}$. We take the center of Δ to be 0 and the radius to be a positive number r, perhaps ∞. We consider complex valued functions defined on Δ. In fact, we concentrate upon functions defined by power series'.

29° Let p_0, p_1, and p_2 be any power series' convergent in Δ, subject to the condition that, for each z in Δ, $p_2(z) \neq 0$. In certain significant contexts, soon to follow, we will describe the power series' f, convergent in Δ, for which:

$$p_2(z)f^{\circ\circ}(z) + p_1(z)f^\circ(z) + p_0(z)f(z) = 0$$

Just as well:

$$(\circ\circ) \qquad f^{\circ\circ}(z) + \frac{p_1(z)}{p_2(z)}f^\circ(z) + \frac{p_0(z)}{p_2(z)}f(z) = 0$$

Constant Coefficients

30° Let us consider first the case of Constant Coefficients, in which the given power series' p_0, p_1, and p_2 are constants. Of course, $p_2 \neq 0$. The equation $(\circ\circ)$ now takes the following form:

$$(C) \qquad p_2 f^{\circ\circ}(z) + p_1 f^\circ z) + p_0 f(z) = 0$$

31° Let ζ be any complex number. For a solution f in $\mathbf{G}$, we propose the following power series:

$$f(z) = exp(\zeta z)$$

Clearly:

$$p_2 f^{\circ\circ}(z) + p_1 f^\circ z) + p_0 f(z) = (p_2\zeta^2 + p_1\zeta + p_0)exp(\zeta z)$$

Hence, f is a solution in $\mathbf{G}$ iff ζ is a zero of the following Quadratic Equation:

$$p_2 z^2 + p_1 z + p_0 = 0$$

The Quadratic Formula yields the following zeros:

$$\zeta_1 = \frac{1}{2p_2}(-p - \sqrt{(p_1^2 - 4p_0 p_2)}\,), \quad \zeta_2 = \frac{1}{2p_2}(-p + \sqrt{(p_1^2 - 4p_0 p_2)}\,)$$

If $\zeta_1 \neq \zeta_2$ then we put forward the following solutions in $\mathbf{G}$:

$$h_1(z) = exp(\zeta_1 z), \quad h_2(z) = exp(\zeta_2 z)$$

The wronskian for h_1 and h_2 is:

$$(\zeta_2 - \zeta_1) exp((\zeta_1 + \zeta_2) z)$$

Hence, h_1 and h_2 compose a basis for $\mathbf{G}$.

32° If $\zeta_1 = \zeta_2$ then we put forward the following solutions in $\mathbf{G}$:

$$h_1(z) = exp(\zeta z), \quad h_2(z) = z\, exp(\zeta z)$$

where $\zeta = \zeta_1 = \zeta_2$. The wronskian proves to be:

$$exp(2\zeta z) \neq 0$$

Hence, h_1 and h_2 compose a basis for $\mathbf{G}$.

33° For the special case:

$$(C) \qquad f^{\circ\circ}(z) + \omega^2 f(z) = 0 \qquad (0 < \omega)$$

we find the following basis of solutions in $\mathbf{G}$:

$$f_1(z) = exp(-i\omega z), \quad f_2(z) = exp(+i\omega z)$$

Hermite

34° Now let us consider the case of Hermite:

$$(H) \qquad f^{\circ\circ}(z) - 2z\, f^{\circ}(z) + \lambda\, f(z) = 0$$

where λ is any complex number. The solutions figure in the Quantum Theory of Simple Harmonic Motion. For a solution f to (H), we propose a power series with undetermined coefficients:

$$f(z) = \sum_{j=0}^{\infty} c_j z^j$$

We find that:

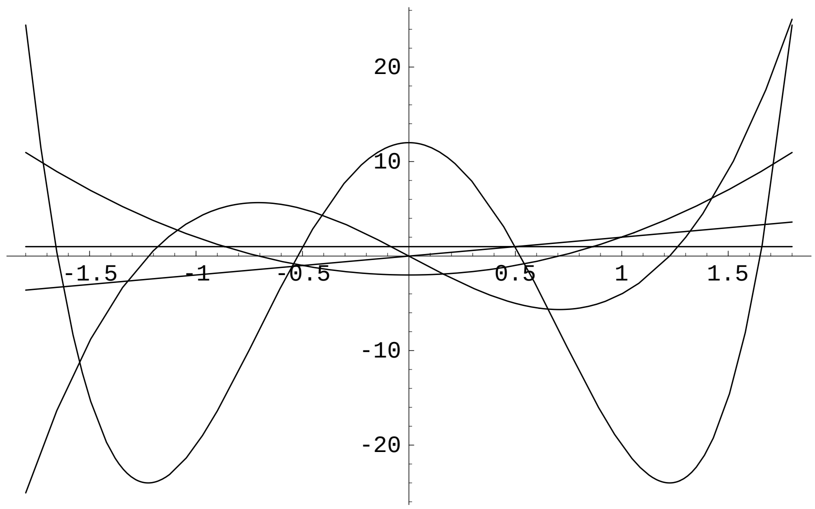
$$\begin{aligned}
f^{\circ\circ}(z) - &2z f^{\circ}(z) + \lambda f(z) \\
&= \sum_{j=2}^{\infty} j(j-1)c_j z^{j-2} - 2z \sum_{j=1}^{\infty} j c_j z^{j-1} + \lambda \sum_{j=0}^{\infty} c_j z^j \\
&= \sum_{k=0}^{\infty} (k+2)(k+1)c_{k+2} z^k - \sum_{k=0}^{\infty} 2k c_k z^k + \sum_{j=0}^{\infty} \lambda c_k z^k \\
&= \sum_{k=0}^{\infty} \Big[(k+2)(k+1)c_{k+2} - (2k-\lambda)c_k\Big] z^k
\end{aligned}$$

Hence, f is a solution to (H) iff the following recursion relation holds:

$$c_{k+2} = \frac{1}{(k+1)(k+2)}(2k-\lambda)c_k$$

where k is any nonnegative integer. Obviously, one can specify the initial coefficients c_0 and c_1 arbitrarily. The rest are then determined. By the Ratio Test, one finds that the radius of convergence for the corresponding power series f is ∞.

35° Let $\lambda = 2\ell$, where ℓ is a nonnegative integer. One can show that there is a polynomial H_ℓ, within constant multiple uniquely defined, which is a solution to (H). Note that the degree of H_ℓ is ℓ. Note that if ℓ is odd then H_ℓ is odd, while if ℓ is even then H_ℓ is even. One refers to H_ℓ as the Hermite Polynomial of degree ℓ. Let us display H_0, H_1, H_2, H_3, and H_4:



Hermite

Legendre

36° In turn, let us consider the case of Legendre:

$$(L') \qquad (1 - z^2)f^{\circ\circ}(z) - 2z\, f^{\circ}(z) + \lambda\, f(z) = 0$$

where λ is any complex number. The solutions figure in the Equation of Laplace and the theory of Spherical Harmonics. For a solution f to (L'), we propose a power series with undetermined coefficients:

$$f(z) = \sum_{j=0}^{\infty} c_j z^j$$

We find that:

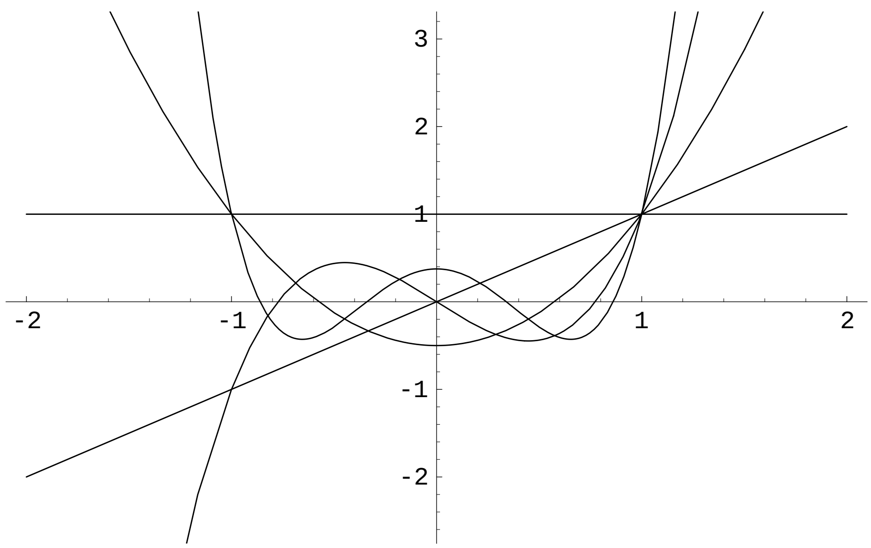
$$\begin{aligned}
&(1 - z^2)f^{\circ\circ}(z) - 2z\, f^{\circ}(z) + \lambda\, f(z) \\
&\quad = (1 - z^2) \sum_{j=2}^{\infty} j(j-1)c_j z^{j-2} - 2z \sum_{j=1}^{\infty} j c_j z^{j-1} + \lambda \sum_{j=0}^{\infty} c_j z^j \\
&\quad = \sum_{k=0}^{\infty} (k+2)(k+1)c_{k+2} z^k - \sum_{k=0}^{\infty} k(k-1)c_k z^k - \sum_{k=0}^{\infty} 2k c_k z^k + \sum_{j=0}^{\infty} \lambda c_k z^k \\
&\quad = \sum_{k=0}^{\infty} \Big[(k+2)(k+1)c_{k+2} - (k(k+1) - \lambda)c_k\Big] z^k
\end{aligned}$$

Hence, f is a solution to (L') iff the following recursion relation holds:

$$c_{k+2} = \frac{1}{(k+1)(k+2)}(k(k+1) - \lambda)c_k$$

where k is any nonnegative integer. Obviously, one can specify the initial coefficients c_0 and c_1 arbitrarily. The rest are then determined. By the Ratio Test, one finds that, typically, the radius of convergence for the corresponding power series f is 1.

37° Let $\lambda = \ell(\ell+1)$, where ℓ is a nonnegative integer. One can show that there is a polynomial L'_ℓ, within constant multiple uniquely defined, which is a solution to (L'). Note that the degree of L'_ℓ is ℓ. Note that if ℓ is odd then L'_ℓ is odd, while if ℓ is even then L'_ℓ is even. One refers to L'_ℓ as the Legendre Polynomial of degree ℓ. Let us display L'_0, L'_1, L'_2, L'_3, and L'_4:



Legendre

Laguerre

38° Let us consider the case of Laguerre:

(L'') $$zf^{\circ\circ}(z) + (1-z)f^{\circ}(z) + \lambda f(z) = 0$$

where λ is any complex number. The solutions figure in the Quantum Theory of the Hydrogen Atom. For a solution f to (L''), we propose a power series with undetermined coefficients:

$$f(z) = \sum_{j=0}^{\infty} c_j z^j$$

We find that:

$$\begin{aligned}
zf^{\circ\circ}(z) + &(1-z)f^{\circ}(z) + \lambda f(z) \\
&= z\sum_{j=2}^{\infty} j(j-1)c_j z^{j-2} + (1-z)\sum_{j=1}^{\infty} jc_j z^{j-1} + \lambda \sum_{j=0}^{\infty} c_j z^j \\
&= \sum_{k=0}^{\infty}(k+1)k\,c_{k+1}z^k + \sum_{k=0}^{\infty}(k+1)c_{k+1}z^k - \sum_{k=0}^{\infty} kc_k z^k + \sum_{j=0}^{\infty} \lambda c_k z^k \\
&= \sum_{k=0}^{\infty}\Big[(k+1)^2 c_{k+1} - (k-\lambda)c_k\Big]z^k
\end{aligned}$$

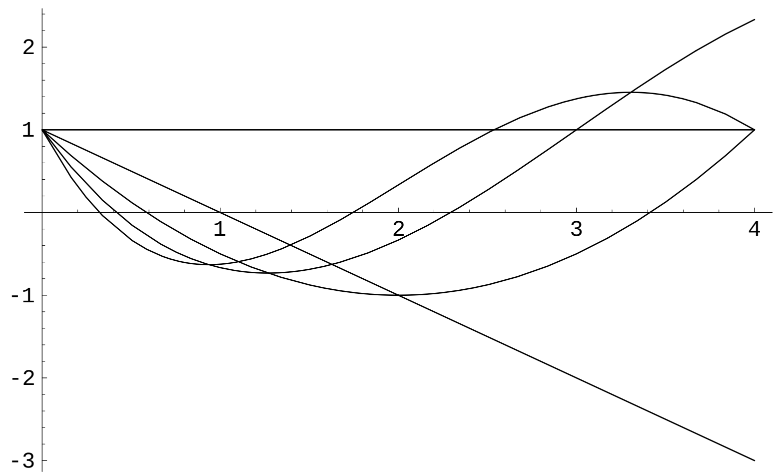
Hence, f is a solution to (L'') iff the following recursion relation holds:

$$c_{k+1} = \frac{1}{(k+1)^2}(k-\lambda)c_k$$

where k is any nonnegative integer. In this notable case, one can freely specify the initial coefficient c_0 but no other. The rest are then determined. By the Ratio Test, one finds that the radius of convergence for the corresponding power series f is ∞.

39° We hasten to note that one can design another solution to (L'') by other methods.

40° Let $\lambda = \ell$, where ℓ is a nonnegative integer. One can show that the solution L''_ℓ to (L''), within constant multiple uniquely defined, is a polynomial. Note that the degree of L''_ℓ is ℓ. One refers to L''_ℓ as the Laguerre Polynomial of degree ℓ. Let us display L''_0, L''_1, L''_2, L''_3, and L''_4:



Laguerre

Bessel

41° Finally, let us consider the celebrated case of Bessel:

$$(B) \qquad z^2 f^{\circ\circ}(z) + z\, f^{\circ}(z) + (z^2 - \lambda^2) f(z) = 0$$

where λ is any complex number. The solutions figure in the theory of the Wave Equation and in many other contexts. Let u and v be the real and imaginary parts of λ:

$$\lambda = u + i\, v$$

Of course, we may assume that:

$$(0 \leq u) \text{ and } ((u = 0) \Longrightarrow (0 \leq v))$$

For a solution f to (B), we propose a function of the following peculiar form:

$$(\circ) \qquad f(z) = z^{\epsilon} \sum_{j=0}^{\infty} c_j z^j$$

where ϵ is any complex number. We intend that:

$$z^{\epsilon} = exp(\epsilon\, log(z))$$

and that z be restricted to the principal domain F of the logarithm function. Without loss of generality, we may assume that:

$$c_0 \neq 0$$

By a pattern of computation now familiar, we find that:

$$\begin{aligned} & z^2 f^{\circ\circ}(z) + z\,f^{\circ}(z) + (z^2 - \lambda^2) f(z) \\ & \quad = z^{\epsilon}\left\{(\epsilon^2 - \lambda^2)c_0 + ((1+\epsilon)^2 - \lambda^2)c_1 z + \sum_{k=2}^{\infty}\big[((k+\epsilon)^2 - \lambda^2)c_k + c_{k-2}\big]z^k\right\} \end{aligned}$$

Hence, f is a solution to (B) iff the following relations hold:

$$(\bullet) \qquad \begin{aligned} (\epsilon^2 - \lambda^2)c_0 &= 0 \\ ((1+\epsilon)^2 - \lambda^2)c_1 &= 0 \\ ((k+\epsilon)^2 - \lambda^2)c_k + c_{k-2} &= 0 \end{aligned}$$

where k is any integer for which $2 \leq k$. Obviously:

$$\epsilon = \pm\lambda$$

Let K be the subset of $\mathcal{Z}^+$ consisting of all positive integers k such that:

$$(k+\epsilon)^2 - \lambda^2 = k(k+2\epsilon) = 0$$

Of course, either $K = \emptyset$ or $K \neq \emptyset$. Let us assume first that $K = \emptyset$. In this case, we may select any two nonzero complex numbers c'_0 and c''_0, to obtain the following two solutions $f_\pm$ to (B):

$$f_-(z) = z^{-\lambda}\sum_{k=0}^{\infty} c'_k z^k, \quad f_+(z) = z^{\lambda}\sum_{k=0}^{\infty} c''_k z^k$$

where:

$$\begin{aligned} c'_1 &= \ \ 0 \\ c'_k &= -\frac{1}{k(k-2\lambda)}c'_{k-2} \end{aligned}$$

and:

$$\begin{aligned} c''_1 &= \ \ 0 \\ c''_k &= -\frac{1}{k(k+2\lambda)}c''_{k-2} \end{aligned}$$

and where k is any integer for which $2 \leq k$. Neither f_- nor f_+ is a constant multiple of the other. Now let us consider the alternate case, in which $K \neq \emptyset$. Let ℓ be any member of K. We find that:

$$\epsilon = -\lambda, \quad \ell + \epsilon = \lambda$$

hence, that:

$$2\lambda = \ell$$

Relations $(\bullet)$ force the following chain of equalities:

$$\ldots\ , \ c_{\ell-6} = 0, \ c_{\ell-4} = 0, \ c_{\ell-2} = 0$$

If ℓ is even then the chain terminates at c_0, contradicting our initial condition that $c_0 \neq 0$. If ℓ is odd then the chain terminates at c_1, without conflict. Hence, we may select any nonzero complex number c'_ℓ, to obtain the following solution f_- to (B):

$$f_-(z) = z^{-\lambda} \sum_{k=\ell}^{\infty} c'_k z^k$$

where:

$$\begin{aligned} c'_{\ell+1} &= \quad 0 \\ c'_k &= -\frac{1}{k(k-2\lambda)} c''_{k-2} \end{aligned}$$

where k is any integer for which $\ell + 2 \leq k$. However:

$$\begin{aligned} f_-(z) &= z^{-\lambda+\ell} \sum_{k=\ell}^{\infty} c'_k z^{k-\ell} \\ &= z^{\lambda} \sum_{n=0}^{\infty} c'_{\ell+n} z^n \end{aligned}$$

and:

$$\frac{1}{k(k-2\lambda)} = \frac{1}{n(n+2\lambda)}$$

where $k = \ell + n$. Hence, f_- merely reproduces the solution f_+ to (B) obtained by setting $\epsilon = \lambda$ and selecting a nonzero complex number c''_0:

$$f_+(z) = z^{\lambda} \sum_{k=0}^{\infty} c''_k z^k$$

where:

$$\begin{aligned} c''_1 &= \quad 0 \\ c''_k &= -\frac{1}{k(k+2\lambda)} c''_{k-2} \end{aligned}$$

and where k is any integer for which $2 \leq k$.

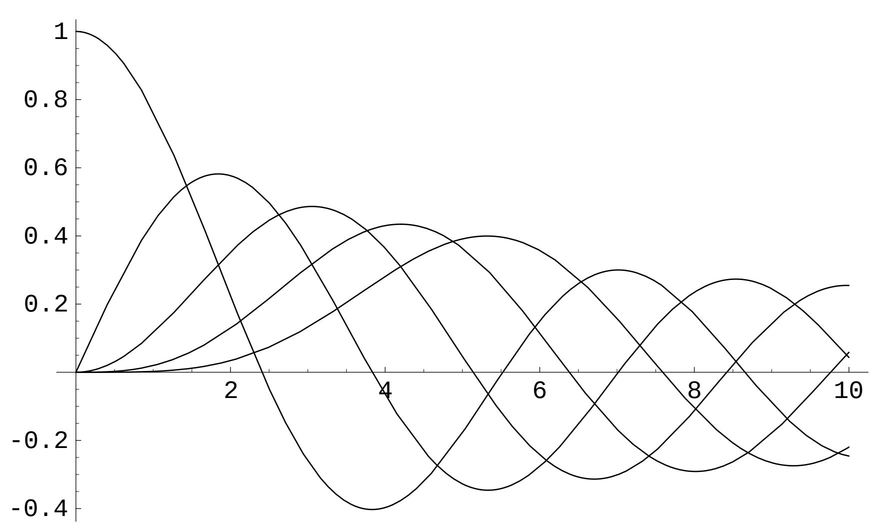
42° One can distinguish the cases in which $K \neq \emptyset$ and $K = \emptyset$ by noting whether or not 2λ is an integer. In the latter case, one obtains two solutions $f_\pm$ to (B) of the form $(\circ)$, neither a constant multiple of the other. In the former case, one obtains just one solution f of the form $(\circ)$. However, one can design another solution by other means.

43° One can show that the radii of convergence for the power series' factors in the various solutions to (B) equal ∞.

44° With reference to the foregoing discussion, let us introduce notation for the Bessel Functions for integral values of λ:

$$J_\lambda : \qquad J_0,\ J_1,\ J_2,\ J_3,\ J_4,\ \ldots$$

For the first few nonnegative integers λ $(0 \leq \lambda \leq 4)$, let us display the graphs of these functions:



Bessel

Chapter 3 STURM/LIOUVILLE THEORY

Introduction

1° In this chapter, we describe one of the cornerstones of modern analysis: the eigenvalue problem for Symmetric Second Order Linear Differential Operators. We consider just the case in which the underlying domain is a closed finite interval. As an application, we obtain a presentation of the theory of Fourier Series.

Second Order Linear Differential Operators

2° Let I be a closed finite interval in $\mathcal{R}$:

$$I \equiv [\,a, b\,] \qquad (a < b)$$

We shall denote by $\mathbf{C}^0$, $\mathbf{C}^1$, and $\mathbf{C}^2$ the (complex) linear spaces consisting of all (complex valued) functions defined on I which are continuous, continuously differentiable, and twice continuously differentiable, respectively.

3° By a Second Order Linear Differentiable Operator on I, one means a (necessarily linear) mapping $\mathbf{L}$ carrying $\mathbf{C}^2$ to $\mathbf{C}^0$ and having the following form:

$$\mathbf{L}(f) = p_2 f^{\circ\circ} + p_1 f^{\circ} + p_0 f \qquad (f \in \mathbf{C}^2)$$

where p_0, p_1, and p_2 are functions in $\mathbf{C}^0$ and where, for each t in I, $p_2(t) \neq 0$. The functions p_0, p_1, and p_2 determine and are determined by $\mathbf{L}$.

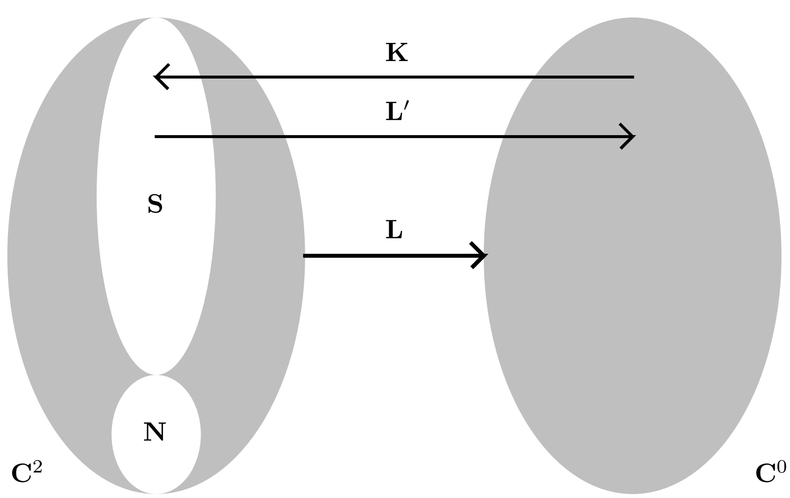
4° Let $\mathbf{L}$ be a Second Order Linear Differentiable Operator on I, determined by the functions p_0, p_1, and p_2, and let $\mathbf{N}$ be the kernel of $\mathbf{L}$. By definition, $\mathbf{N}$ is the linear subspace of $\mathbf{C}^2$ consisting of all functions f for which $\mathbf{L}(f) = 0$. By the general existence/uniqueness theory for Ordinary Differential Equations, $\mathbf{L}$ is surjective and $\mathbf{N}$ is two dimensional.

5° Now let $\mathbf{S}$ be a linear subspace of $\mathbf{C}^2$ and let $\mathbf{L}'$ be the restriction of $\mathbf{L}$ to $\mathbf{S}$. We contend that $\mathbf{L}'$ is bijective iff $\mathbf{N}$ and $\mathbf{S}$ compose a direct sum decomposition of $\mathbf{C}^2$. We mean to say that, for each f in $\mathbf{C}^2$, there exist h in $\mathbf{N}$ and f' in $\mathbf{S}$ such that $f = h + f'$. Moreover, for any h_1 and h_2 in $\mathbf{N}$ and for any f_1' and f_2' in $\mathbf{S}$, if $h_1 + f_1' = h_2 + f_2'$ then $h_1 = h_2$ and $f_1' = f_2'$. The latter condition means, quite simply, that $\mathbf{N} \cap \mathbf{S} = \{0\}$. Together, the conditions mean that every function in $\mathbf{C}^2$ can be expressed uniquely as the sum of a function in $\mathbf{N}$ and a function in $\mathbf{S}$.

6° Let us prove the contention. To that end, let us assume that $\mathbf{L}'$ is bijective. Let f be any function in $\mathbf{C}^2$. Since $\mathbf{L}'$ is surjective, there exists f' in $\mathbf{S}$ such that $\mathbf{L}'(f') = \mathbf{L}(f)$. Let $h \equiv f - f'$. Obviously, h is in $\mathbf{N}$ and $f = h + f'$. In turn, let f be any function in $\mathbf{N} \cap \mathbf{S}$. That is, $\mathbf{L}'(f) = 0$ and $f \in \mathbf{S}$. Since $\mathbf{L}'$ is injective, $f = 0$. We conclude that $\mathbf{N}$ and $\mathbf{S}$ compose a direct sum decomposition of $\mathbf{C}^2$.

7° Now let us assume that $\mathbf{N}$ and $\mathbf{S}$ compose a direct sum decomposition of $\mathbf{C}^2$. For convenience of expression, let us refer to this assumption as Δ. Let g be any function in $\mathbf{C}^0$. Since $\mathbf{L}$ is surjective, there exists f in $\mathbf{C}^2$ such that $\mathbf{L}(f) = g$. By Δ, there exist h in $\mathbf{N}$ and f' in $\mathbf{S}$ such that $f = h + f'$. Obviously, $\mathbf{L}'(f') = g$. Consequently, $\mathbf{L}'$ is surjective. In turn, let f' be a function in $\mathbf{S}$ such that $\mathbf{L}'(f') = 0$. Obviously, $f' \in \mathbf{N} \cap \mathbf{S}$. By Δ, $f' = 0$. Consequently, $\mathbf{L}'$ is injective. We conclude that $\mathbf{L}'$ is bijective.

8° Under the conditions just described, the inverse $\mathbf{K}$ of $\mathbf{L}'$ would carry $\mathbf{C}^0$ bijectively to $\mathbf{S}$. We plan to study a broad class of linear subspaces $\mathbf{S}$, defined by Boundary Conditions, for which $\mathbf{K}$ admits an elegant description as a Linear Integral Operator. See Theorem A. Later, we shall concentrate upon the special cases in which the linear operators $\mathbf{L}$ and linear subspaces $\mathbf{S}$ are Symmetric. In such cases, the inverse $\mathbf{K}$ of $\mathbf{L}'$ admits detailed analysis in terms of eigenfunctions and eigenvalues. See Theorem B.



Differential/Integral Operators

9° To describe the appropriate linear subspaces $\mathbf{S}$ of $\mathbf{C}^2$, we proceed as follows. Let:

$$\sigma_1, \sigma_1^\circ, \tau_1, \tau_1^\circ, \sigma_2, \sigma_2^\circ, \tau_2, \tau_2^\circ$$

be any eight complex numbers. Let $\mathbf{S}$ be the linear subspace of $\mathbf{C}^2$ consisting of all functions f which satisfy the following Boundary Conditions:

$$(BC) \qquad \begin{aligned} \sigma_1 f(a) + \sigma_1^\circ f^\circ(a) + \tau_1 f(b) + \tau_1^\circ f^\circ(b) = 0 \\ \sigma_2 f(a) + \sigma_2^\circ f^\circ(a) + \tau_2 f(b) + \tau_2^\circ f^\circ(b) = 0 \end{aligned}$$

We hasten to express the foregoing conditions in a more efficient form, as follows. Let us assemble the foregoing numbers as a matrix:

$$\begin{pmatrix} \sigma_1 & \sigma_1^\circ & \tau_1 & \tau_1^\circ \\ \sigma_2 & \sigma_2^\circ & \tau_2 & \tau_2^\circ \end{pmatrix}$$

and let us denote by m the corresponding linear mapping carrying $\mathcal{C}^4$ to $\mathcal{C}^2$:

$$m(\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}) = \begin{pmatrix} \sigma_1 & \sigma_1^\circ & \tau_1 & \tau_1^\circ \\ \sigma_2 & \sigma_2^\circ & \tau_2 & \tau_2^\circ \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \qquad (\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \in \mathcal{C}^4)$$

In turn, let ζ be the Evaluation Mapping carrying $\mathbf{C}^2$ to $\mathcal{C}^4$:

$$\zeta(f) = \begin{pmatrix} f(a) \\ f^\circ(a) \\ f(b) \\ f^\circ(b) \end{pmatrix} \qquad (f \in \mathbf{C}^2)$$

Finally, let μ be the composition of m and ζ, carrying $\mathbf{C}^2$ to $\mathcal{C}^2$:

$$\begin{aligned} \mu(f) &= m(\zeta(f)) \\ &= \begin{pmatrix} \sigma_1 & \sigma_1^\circ & \tau_1 & \tau_1^\circ \\ \sigma_2 & \sigma_2^\circ & \tau_2 & \tau_2^\circ \end{pmatrix} \begin{pmatrix} f(a) \\ f^\circ(a) \\ f(b) \\ f^\circ(b) \end{pmatrix} \qquad (f \in \mathbf{C}^2) \end{aligned}$$

Now we can describe $\mathbf{S}$ simply as the kernel of μ. That is, for each f in $\mathbf{C}^2$, $f \in \mathbf{S}$ iff:

$$(BC) \qquad \mu(f) = 0$$

10° Let $\mathbf{S}$ be the linear subspace of $\mathbf{C}^2$ defined by the boundary conditions (BC). We contend that $\mathbf{N}$ and $\mathbf{S}$ compose a direct sum decomposition of $\mathbf{C}^2$ iff, for some (and hence for any) basis $\{h_1, h_2\}$ for $\mathbf{N}$:

$$(MZ) \qquad\qquad det\,(\,\mu(h_1)\ \mu(h_2)\,) \neq 0$$

11° Let us assume that $\mathbf{N}$ and $\mathbf{S}$ compose a direct sum decomposition of $\mathbf{C}^2$. Let $\{h_1, h_2\}$ be any basis for $\mathbf{N}$. Let us suppose that condition (MZ) fails. It would follow that there exist complex numbers c_1 and c_2 such that $|c_1|^2 + |c_2|^2 \neq 0$ and $c_1\mu(h_1) + c_2\mu(h_2) = 0$. That is, there exists a function h in $\mathbf{C}^2$, namely, $h = c_1h_1 + c_2h_2$, such that $h \neq 0$ and $h \in \mathbf{N} \cap \mathbf{S}$, in contradiction with our assumption. Consequently, our supposition must be false. We conclude that condition (MZ) holds.

12° Now let us assume that there is a basis $\{h_1, h_2\}$ for $\mathbf{N}$ such that condition (MZ) holds. Let ν be the mapping carrying $\mathcal{C}^2$ to $\mathbf{C}^2$, defined as follows:

$$\nu(\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}) = c_1h_1 + c_2h_2 \qquad (\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in \mathcal{C}^2)$$

Obviously, the range of ν is $\mathbf{N}$. The composition $\mu \cdot \nu$ of ν and μ carries $\mathcal{C}^2$ to itself. It takes the form:

$$(\mu \cdot \nu)(\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}) = (\,\mu(h_1)\ \mu(h_2)\,)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \qquad (\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in \mathcal{C}^2)$$

Since (MZ) holds, $\mu \cdot \nu$ is bijective. Let f be any function in $\mathbf{C}^2$. Let h be the function $\nu((\mu\cdot\nu)^{-1}(\mu(f)))$. Obviously, $h \in \mathbf{N}$. Moreover, $\mu(f - h) = 0$, so $f - h \in \mathbf{S}$. Of course, $f = h + (f - h)$. In turn, let f be a function in $\mathbf{N} \cap \mathbf{S}$. Let c be the member of $\mathcal{C}^2$ for which $\nu(c) = f$. Of course, $(\mu \cdot \nu)(c) = \mu(f) = 0$. Hence, $c = 0$, so $f = 0$. We conclude that $\mathbf{N}$ and $\mathbf{S}$ compose a direct sum decomposition of $\mathbf{C}^2$.

13° At this point, let us fix in mind:

(•) a Second Order Linear Differential Operator $\mathbf{L}$ on I determined by the functions p_0, p_1, and p_2 in $\mathbf{C}^0$

(•) a linear subspace $\mathbf{S}$ of $\mathbf{C}^2$ defined by the Boundary Conditions (BC), subject to the condition (MZ)

By the foregoing discussion, the restriction $\mathbf{L}'$ of $\mathbf{L}$ to $\mathbf{S}$ is bijective. We plan to analyze the inverse $\mathbf{K}$ of $\mathbf{L}'$.

Linear Integral Operators

14° By a Linear Integral Operator on I, one means a (necessarily linear) mapping $\mathbf{K}$ carrying $\mathbf{C}^0$ to itself and having the following form:

$$\mathbf{K}(g)(t) = \int_a^b K(t,u)g(u)du \qquad (g \in \mathbf{C}^0)$$

where K is a continuous (complex valued) function defined on $I \times I$. One says that K defines $\mathbf{K}$.

15° Let us pause to verify that the foregoing definition makes sense. We must show that, for each g in $\mathbf{C}^0$, $\mathbf{K}(g)$ is in $\mathbf{C}^0$ as well. Let g be any function in $\mathbf{C}^0$. Let t be any number in I. Let ϵ be any positive real number. We claim that there is a positive real number δ such that, for any ordered pair $(\bar{t}, u)$ in $I \times I$, if $|\bar{t} - t| \leq \delta$ then $|K(\bar{t}, u) - K(t,u)| \leq \epsilon$. Let us suppose that the claim is false. Under that supposition, we may introduce a sequence:

$$(\bar{t}_1, u_1), (\bar{t}_2, u_2), \ \ldots, \ (\bar{t}_j, u_j), \ \ldots$$

such that, for each positive integer j:

$$|\bar{t}_j - t| \leq \frac{1}{j} \ \text{ and } \ \epsilon < |K(\bar{t}_j, u_j) - K(t, u_j)|$$

By the Bolzano/Weierstrass Theorem, we may presume that the sequence is convergent:

$$(\bar{t}_j, u_j) \longrightarrow (t, \bar{u})$$

Of course, K is continuous at $(t, \bar{u})$. Consequently:

$$\epsilon \leq |K(t, \bar{u}) - K(t, \bar{u})|$$

a contradiction. We infer that the supposition is false.

16° Consequently, the claim is true. Let us introduce a positive number δ of the sort described. Let $\|g\|$ be the uniform norm for g on I:

$$\|g\| \equiv sup\{\, |g(u)| \, : \, a \leq u \leq b \,\}$$

Now we find that, for each $\bar{t}$ in I, if $|\bar{t} - t| \leq \delta$ then:

$$|\mathbf{K}(g)(\bar{t}) - \mathbf{K}(g)(t)| \leq \int_a^b |K(\bar{t}, u) - K(t,u)||g(u)|du \leq \epsilon \|g\|(b-a)$$

We infer that $\mathbf{K}(g)$ is continuous at t. Hence, $\mathbf{K}(g)$ is in $\mathbf{C}^0$.

17° We contend that there is precisely one K such that $\mathbf{K}$ carries $\mathbf{C}^0$ to $\mathbf{S}$ and such that $\mathbf{L}'$ and $\mathbf{K}$ are inverse to one another. Consequently, given a function g in $\mathbf{C}^0$, we may (in principle) solve the differential equation:

$$\mathbf{L}(?) = g$$

by integration:

$$f = \mathbf{K}(g)$$

Moreover, the solution f satisfies the stated boundary conditions. Subject to those conditions, it is unique.

18° One refers to K as the Green Function for $\mathbf{L}$ and $\mathbf{S}$.

19° Let us prove the contention. To that end, we introduce a basis $\{h_1, h_2\}$ for $\mathbf{N}$. Since the wronskian for $\{h_1, h_2\}$ never vanishes on I, we may introduce functions θ_1 and θ_2 in $\mathbf{C}^0$ such that:

$$\begin{pmatrix} h_1 & h_2 \\ h_1^\circ & h_2^\circ \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ p_2^{-1} \end{pmatrix} \tag{1}$$

In turn, let us introduce functions v_1, v_2, w_1, and w_2 in $\mathbf{C}^0$ subject to the conditions:

$$\begin{aligned} \theta_1 &= w_1 - v_1 \\ \theta_2 &= w_2 - v_2 \end{aligned} \tag{2}$$

For now, we require nothing more. Finally, let us define the function K:

$$K(t,u) = \begin{cases} h_1(t)\, v_1(u) + h_2(t)\, v_2(u) & \text{if } t \le u \\ h_1(t) w_1(u) + h_2(t) w_2(u) & \text{if } u \le t \end{cases} \qquad ((t,u) \in I \times I)$$

By relations (1) and (2):

$$h_1(w_1 - v_1) + h_2(w_2 - v_2) = 0$$

Hence, the definition of K on the diagonal of $I \times I$ is unambiguous. Moreover, K is continuous on $I \times I$.

20° Let g be any function in $\mathbf{C}^0$ and let f be $\mathbf{K}(g)$:

$$f(t) = \int_a^b K(t,u) g(u) du \qquad (t \in I)$$

21° We claim that $\mathbf{L}(f) = g$. To prove the claim, let us introduce the functions:

$$\text{(3)} \qquad \begin{aligned} \gamma_1(t) &= \int_a^t w_1(u)g(u)du + \int_t^b v_1(u)g(u)du \\ \gamma_2(t) &= \int_a^t w_2(u)g(u)du + \int_t^b v_2(u)g(u)du \end{aligned} \qquad (t \in I)$$

By straightforward calculation from the definition:

$$\text{(4)} \qquad f(t) = (\, h_1(t) \;\, h_2(t)\,) \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix}$$

By relation (2):

$$\text{(5)} \qquad \begin{aligned} \gamma_1^\circ(t) &= \theta_1(t)g(t) \\ \gamma_2^\circ(t) &= \theta_2(t)g(t) \end{aligned}$$

By relations (1), (4), and (5):

$$\text{(6)} \qquad \begin{aligned} f^\circ(t) &= (\, h_1^\circ(t) \;\, h_2^\circ(t)\,) \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} + (\, h_1(t) \;\, h_2(t)\,) \begin{pmatrix} \gamma_1^\circ(t) \\ \gamma_2^\circ(t) \end{pmatrix} \\ &= (\, h_1^\circ(t) \;\, h_2^\circ(t)\,) \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} \end{aligned}$$

By relations (1), (5), and (6):

$$\text{(7)} \qquad \begin{aligned} f^{\circ\circ}(t) &= (\, h_1^{\circ\circ}(t) \;\, h_2^{\circ\circ}(t)\,) \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} + (\, h_1^\circ(t) \;\, h_2^\circ(t)\,) \begin{pmatrix} \gamma_1^\circ(t) \\ \gamma_2^\circ(t) \end{pmatrix} \\ &= (\, h_1^{\circ\circ}(t) \;\, h_2^{\circ\circ}(t)\,) \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} + p_2^{-1}(t)g(t) \end{aligned}$$

We note in passing that f must lie in $\mathbf{C}^2$. Assembling relations (4), (6), and (7), we obtain:

$$\begin{aligned} \mathbf{L}(f) &= p_0 f + p_2 f^\circ + p_2 f^{\circ\circ} \\ &= (\, \mathbf{L}(h_1) \;\, \mathbf{L}(h_2)\,) \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} + p_2 p_2^{-1} g \\ &= g \end{aligned}$$

22° We are pleased to find that $\mathbf{L}(f) = g$ but, at this point, we cannot show that $f \in \mathbf{S}$. To do so, we must refine our description of the functions v_1, v_2, w_1, and w_2.

23° From relations (3), we find that:

$$\begin{aligned} \gamma_1(a) &= \int_a^b v_1(u)g(u)du \\ \gamma_2(a) &= \int_a^b v_2(u)g(u)du \\ \gamma_1(b) &= \int_a^b w_1(u)g(u)du \\ \gamma_2(b) &= \int_a^b w_2(u)g(u)du \end{aligned} \tag{8}$$

By relations (4) and (6):

$$\begin{pmatrix} f(a) \\ f^\circ(a) \\ f(b) \\ f^\circ(b) \end{pmatrix} = \begin{pmatrix} h_1(a) & h_2(a) & 0 & 0 \\ h_1^\circ(a) & h_2^\circ(a) & 0 & 0 \\ 0 & 0 & h_1(b) & h_2(b) \\ 0 & 0 & h_1^\circ(b) & h_2^\circ(b) \end{pmatrix} \begin{pmatrix} \gamma_1(a) \\ \gamma_2(a) \\ \gamma_1(b) \\ \gamma_2(b) \end{pmatrix} \tag{9}$$

By inspecting relations (8), we see that if (!):

$$\begin{pmatrix} \sigma_1 & \sigma_1^\circ & \tau_1 & \tau_1^\circ \\ \sigma_2 & \sigma_2^\circ & \tau_2 & \tau_2^\circ \end{pmatrix} \begin{pmatrix} h_1(a) & h_2(a) & 0 & 0 \\ h_1^\circ(a) & h_2^\circ(a) & 0 & 0 \\ 0 & 0 & h_1(b) & h_2(b) \\ 0 & 0 & h_1^\circ(b) & h_2^\circ(b) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{10}$$

then $f \in \mathbf{S}$.

24° Let us proceed to describe functions v_1, v_2, w_1, and w_2 which meet condition (10) while preserving the relations (2) described earlier:

$$\begin{aligned} v_1 &= w_1 - \theta_1 \\ v_2 &= w_2 - \theta_2 \end{aligned} \tag{2}$$

To that end, we apply relations (2) to reform relation (10):

$$\begin{pmatrix} \sigma_1 & \sigma_1^\circ & \tau_1 & \tau_1^\circ \\ \sigma_2 & \sigma_2^\circ & \tau_2 & \tau_2^\circ \end{pmatrix} \begin{pmatrix} h_1(a) & h_2(a) & 0 & 0 \\ h_1^\circ(a) & h_2^\circ(a) & 0 & 0 \\ 0 & 0 & h_1(b) & h_2(b) \\ 0 & 0 & h_1^\circ(b) & h_2^\circ(b) \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \tag{11}$$

where:

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \equiv \begin{pmatrix} \sigma_1 & \sigma_1^\circ & \tau_1 & \tau_1^\circ \\ \sigma_2 & \sigma_2^\circ & \tau_2 & \tau_2^\circ \end{pmatrix} \begin{pmatrix} h_1(a) & h_2(a) & 0 & 0 \\ h_1^\circ(a) & h_2^\circ(a) & 0 & 0 \\ 0 & 0 & h_1(b) & h_2(b) \\ 0 & 0 & h_1^\circ(b) & h_2^\circ(b) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ 0 \\ 0 \end{pmatrix}$$

By careful inspection, we find that relation (11) collapses to the following simple form:

$$(12) \qquad (\mu(h_1) \ \mu(h_2)) \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

By condition (MZ), the matrix:

$$(\mu(h_1) \ \mu(h_2))$$

is invertible. Now it is plain that we can produce functions v_1, v_2, w_1, and w_2 with the required properties.

25° Finally, we must show that the Green Function K is unique. To that end, let us introduce functions K_1 and K_2 such that both K_1 and K_2 define $\mathbf{K}$. We must show that $K_1 = K_2$.

26° Let K_1^r and K_2^r be the real parts and let K_1^i and K_2^i be the imaginary parts of K_1 and K_2, respectively. Of course, we must show that $K_1^r = K_2^r$ and $K_1^i = K_2^i$. Let us suppose that $K_1^r \neq K_2^r$. Let (t, v) be a member of $I \times I$ such that $K_1^r(t, v) \neq K_2^r(t, v)$. Without loss of generality, we might as well suppose that $K_1^r(t, v) < K_2^r(t, v)$. Under this supposition, we may introduce a positive number δ such that, for each u in I, if $|u - v| \leq \delta$ then $K_1^r(t, u) < K_2^r(t, u)$. Now let us design a function g in $\mathbf{C}^0$ such that g is real valued and such that, for each u in I:

$$\begin{cases} g(u) = 0 & \text{if } u \leq v - \delta \\ 0 < g(u) & \text{if } v - \delta < u < v + \delta \\ g(u) = 0 & \text{if } v + \delta \leq u \end{cases}$$

Let $\mathbf{K}(g)(t)^r$ be the real part of $\mathbf{K}(g)(t)$. We find that:

$$\mathbf{K}(g)(t)^r = \int_a^b K_1^r(t, u) g(u) du < \int_a^b K_2^r(t, u) g(u) du = \mathbf{K}(g)(t)^r$$

By this contradiction, we infer that our supposition is false. It follows that $K_1^r = K_2^r$.

27° By a similar argument, one may show that $K_1^i = K_2^i$. We conclude that $K_1 = K_2$. •

The Symmetric Case

28° Let us supply $\mathbf{C}^0$ with a Geometric Structure, by means of the following Inner Product:

$$\langle\!\langle\, g_1, g_2 \,\rangle\!\rangle \equiv \frac{1}{d}\int_a^b g_1(t)\bar{g}_2(t)dt \qquad (g_1, g_2 \in \mathbf{C}^0,\ \ d \equiv (b-a))$$

and the corresponding Integral Norm:

$$\langle\!\langle g \rangle\!\rangle \equiv \sqrt{\langle\!\langle\, g, g \,\rangle\!\rangle} \qquad (g \in \mathbf{C}^0)$$

The bar signals complex conjugation. We plan to introduce conditions on $\mathbf{L}$ and $\mathbf{S}$ under which the operator $\mathbf{L}'$ is Symmetric, in the sense that:

$$(SA) \qquad \langle\!\langle\, \mathbf{L}(f_1), f_2 \,\rangle\!\rangle = \langle\!\langle\, f_1, \mathbf{L}(f_2) \,\rangle\!\rangle \qquad (f_1, f_2 \in \mathbf{S})$$

To that end, we assume that:

(1) p_0, p_1, and p_2 are real valued
(2) $p_1 \in \mathbf{C}^1$ and $p_2 \in \mathbf{C}^2$
(3) $p_1 = p_2^\circ$

Granted conditions (1), (2), and (3), we can easily verify the Identity of Lagrange:

$$(LL) \qquad \mathbf{L}(f_1)\,\bar{f}_2 - f_1\,\overline{\mathbf{L}(f_2)} = B(f_1, f_2)^\circ \qquad (f_1, f_2 \in \mathbf{C}^2)$$

where:

$$(BB) \qquad B(f_1, f_2) = p_2(f_1^\circ \bar{f}_2 - f_1 \bar{f}_2^\circ)$$

In turn, we assume that:

(4) for any f_1 and f_2 in $\mathbf{S}$, $B(f_1, f_2)\big|_a^b = 0$

Now we find that, under conditions (1), (2), (3), and (4), $\mathbf{L}'$ is Symmetric:

$$\begin{aligned}
\langle\!\langle\, \mathbf{L}(f_1), f_2 \,\rangle\!\rangle - \langle\!\langle\, f_1, \mathbf{L}(f_2) \,\rangle\!\rangle &= \frac{1}{d}\int_a^b [\mathbf{L}(f_1)(t)\,\bar{f}_2(t) - f_1(t)\,\overline{\mathbf{L}(f_2)}(t)]dt \\
&= \frac{1}{d}\int_a^b B(f_1, f_2)^\circ(t)dt \\
&= \frac{1}{d}\,B(f_1, f_2)\big|_a^b \\
&= 0
\end{aligned}$$

29° We must confess that, while conditions (1), (2), (3), and (4) are sufficient to imply that $\mathbf{L}'$ is Symmetric, they are far from necessary. However, a full development of the matter would require a substantial digression.

An Illustration

30° For an illustration, let us take the interval I to be $[-\pi, \pi]$, let us define the operator $\mathbf{L}$ by the relation:

$$\mathbf{L}(f) = -f^{\circ\circ} + f \qquad (f \in \mathbf{C}^2)$$

and let us set the boundary conditions:

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} f(-\pi) \\ f^{\circ}(-\pi) \\ f(\pi) \\ f^{\circ}(\pi) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

That is:

$$f \in \mathbf{S} \quad \text{iff} \quad f(-\pi) = f(\pi) \ \text{ and } \ f^{\circ}(-\pi) = f^{\circ}(\pi)$$

Obviously, conditions (1), (2), and (3) hold. We introduce the basis:

$$h_1(t) = exp(-t), \;\; h_2(t) = exp(t), \qquad (-\pi \leq t \leq \pi)$$

for $\mathbf{N}$ and we verify that condition (MZ) holds, as follows:

$$\begin{aligned} det\left[\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} h_1(-\pi) & h_2(-\pi) \\ h_1^{\circ}(-\pi) & h_2^{\circ}(-\pi) \\ h_1(\pi) & h_2(\pi) \\ h_2^{\circ}(\pi) & h_2^{\circ}(\pi) \end{pmatrix}\right] \\ = det\begin{pmatrix} exp(\pi) - exp(-\pi) & exp(-\pi) - exp(\pi) \\ -exp(\pi) + exp(-\pi) & exp(-\pi) - exp(\pi) \end{pmatrix} \\ \neq 0 \end{aligned}$$

In turn, we check that condition (4) holds, as follows:

$$\begin{aligned} & B(f_1, f_2)\big|_{-\pi}^{\pi} \\ & \quad = p_2(f_1^{\circ} f_2 - f_1 \bar{f}_2^{\circ})_{-\pi}^{\pi} \\ & \quad = -\big(f_1^{\circ}(\pi) f_2(\pi) - f_1(\pi) \bar{f}_2^{\circ}(\pi) - f_1^{\circ}(-\pi) f_2(-\pi) + f_1(-\pi) \bar{f}_2^{\circ}(-\pi)\big) \\ & \quad = 0 \end{aligned}$$

where f_1 and f_2 are any functions in $\mathbf{S}$.

Eigenvalues and Eigenfunctions

31° Let us return to the operator $\mathbf{L}$ and the subspace $\mathbf{S}$, subject as usual to conditions (BC) and (MZ) but subject to conditions (1), (2), (3), and (4) as well. Let $\mathbf{L}'$ be the restriction of $\mathbf{L}$ to $\mathbf{S}$. Of course, $\mathbf{L}'$ carries $\mathbf{S}$ bijectively to $\mathbf{C}^0$ and $\mathbf{L}'$ is Symmetric.

32° We plan to analyze $\mathbf{L}'$ in terms of the concepts of Eigenvalue and Eigenfunction. Let λ be a complex number. Let $\mathbf{\Lambda}$ be the subset of $\mathbf{S}$ consisting of all functions f for which:

$$(EV) \qquad\qquad \mathbf{L}(f) = \lambda f$$

Clearly, $\mathbf{\Lambda}$ is a linear subspace of $\mathbf{S}$. By the Fundamental Theorem, the dimension of $\mathbf{\Lambda}$ must be 0, 1, or 2. When the dimension of $\mathbf{\Lambda}$ is at least 1, we refer to λ as an Eigenvalue for $\mathbf{L}'$, to $\mathbf{\Lambda}$ itself as the Eigenspace for $\mathbf{L}'$ relative to λ, and to the various functions f in $\mathbf{\Lambda}$ as the Eigenfunctions for $\mathbf{L}'$ relative to λ.

33° By relation (EV):

$$\lambda\langle\!\langle f, f \rangle\!\rangle = \langle\!\langle \lambda f, f \rangle\!\rangle = \langle\!\langle \mathbf{L}(f), f \rangle\!\rangle = \langle\!\langle f, \mathbf{L}(f) \rangle\!\rangle = \langle\!\langle f, \lambda f \rangle\!\rangle = \bar{\lambda}\langle\!\langle f, f \rangle\!\rangle$$

Consequently, eigenvalues for $\mathbf{L}'$ must be real numbers.

The Illustration: Redux

34° In context of the foregoing illustration, we seek solutions of the relation:

$$\mathbf{L}(f) = -f^{\circ\circ} + f = \lambda f$$

subject to the boundary conditions:

$$f(-\pi) = f(\pi) \text{ and } f^{\circ}(-\pi) = f^{\circ}(\pi)$$

We find:

$$f(t) = c_1 exp(-ikt) + c_2 exp(ikt) \qquad (-\pi \leq t \leq \pi)$$

where c_1 and c_2 are any complex numbers and where k is any integer for which $k^2 = \lambda - 1$. Consequently, the eigenvalues λ stand as follows:

$$1, 2, 5, 10, \ \ldots$$

The dimensions of the corresponding eigenspaces $\mathbf{\Lambda}$ are 1 for $\lambda = 1$ and 2 for the rest.

Theorem B

35° We contend that there are a sequence:

$$\Lambda : \qquad \lambda_1, \lambda_2, \lambda_3, \ \dots \ , \lambda_j, \ \dots$$

of (nonzero) real numbers and a sequence:

$$H : \qquad h_1, h_2, h_3, \ \dots \ , h_j, \ \dots$$

of functions in $\mathbf{S}$ such that:

($B1$) for each index j, $\mathbf{L}(h_j) = \lambda_j h_j$
($B2$) for any indices k and ℓ:

$$\langle\!\langle\, h_k, h_\ell \,\rangle\!\rangle = \begin{cases} 0 & \text{if } k \neq \ell \\ 1 & \text{if } k = \ell \end{cases}$$

($B3$) for each index j, $|\lambda_j| \leq |\lambda_{j+1}|$; moreover, $|\lambda_j| \uparrow \infty$
($B4$) for each function f in $\mathbf{S}$, the series:

$$\sum_{j=1}^{\infty} \langle\!\langle\, f, h_j \,\rangle\!\rangle h_j$$

converges uniformly to f.

36° Before turning to the proof of the theorem, let us emphasize certain features of the foregoing assertions. By ($B2$), the terms of H compose an Orthonormal Family in $\mathbf{S}$. By ($B1$), the terms of Λ are eigenvalues for $\mathbf{L}'$ and the terms of H are corresponding eigenfunctions. Since $\mathbf{L}'$ is injective, the terms of Λ must be nonzero. By ($B3$), the absolute values of the terms of Λ increase without limit. However, certain terms may occur twice, namely, those for which the corresponding eigenspace has dimension 2.

37° Regarding ($B4$), let us just say that we will explain the sense of it very carefully in the following articles.

The Green Function

38° For the proof of the theorem, we direct our attention to the inverse $\mathbf{K}$ of $\mathbf{L}'$. We do so because the analysis of integral operators is substantially simpler than the analysis of differential operators.

39° Let K be the Green Function for $\mathbf{L}$ and $\mathbf{S}$. By definition, K defines $\mathbf{K}$. In our present context, it is plain that K is real valued. Moreover, the condition (SA) for $\mathbf{L}'$ carries over to $\mathbf{K}$:

$$(SA) \qquad \langle\!\langle\, \mathbf{K}(g_1), g_2 \,\rangle\!\rangle = \langle\!\langle\, g_1, \mathbf{K}(g_2) \,\rangle\!\rangle \qquad (g_1, g_2 \in \mathbf{C}^0)$$

40° For the analysis of $\mathbf{K}$, we require not only the Integral Norm but also the Uniform Norm. For convenience, let us display them together:

$$\begin{aligned} \langle\!\langle g \rangle\!\rangle &= \sqrt{\frac{1}{d}\int_a^b |g(t)|^2 dt} \\ \|g\| &= sup\,\{\, |g(t)| \,:\, a \leq t \leq b \,\} \end{aligned} \qquad (g \in \mathbf{C}^0)$$

Obviously:

$$(13) \qquad \langle\!\langle g \rangle\!\rangle \leq \|g\|$$

41° Let $\|K\|$ be the uniform norm for K itself:

$$\|K\| = sup\,\{\, |K(t,u)| \,:\, (t,u) \in I \times I \,\}$$

We contend that, for each function g in $\mathbf{C}^0$:

$$(14) \qquad \|\mathbf{K}(g)\| \leq d\,\|K\|\langle\!\langle g \rangle\!\rangle$$

Let t be any number in I. To prove the contention, we must prove that:

$$|\mathbf{K}(g)(t)| \leq d\,\|K\|\langle\!\langle g \rangle\!\rangle$$

To that end, let K_t be the function in $\mathbf{C}^0$ defined as follows:

$$K_t(u) \equiv K(t,u) \qquad (u \in I)$$

Obviously:

$$\frac{1}{d}\mathbf{K}(g)(t) = \langle\!\langle\, K_t, \bar{g} \,\rangle\!\rangle$$

By the Cauchy/Schwarz Inequality and by relation (13), we find that:

$$\frac{1}{d}|\mathbf{K}(g)(t)| = |\langle\!\langle\, K_t, \bar{g} \,\rangle\!\rangle| \leq \langle\!\langle K_t \rangle\!\rangle\langle\!\langle \bar{g} \rangle\!\rangle \leq \|K_t\|\langle\!\langle \bar{g} \rangle\!\rangle \leq \|K\|\langle\!\langle g \rangle\!\rangle$$

42° Relation (13) asserts that the uniform norm dominates the integral norm but relation (14) asserts that, after application of the integral operator $\mathbf{K}$, the integral norm dominates the uniform norm.

43° Now we contend that, for each function g in $\mathbf{C}^0$ and for each positive number ϵ, there is a positive number δ such that, for any numbers t and $\bar{t}$ in I, if $|\bar{t} - t| \leq \delta$ then:

$$|\mathbf{K}(g)(\bar{t}) - \mathbf{K}(g)(t)| \leq d\,\epsilon \langle\!\langle g \rangle\!\rangle \tag{15}$$

To prove the contention, we need only review article 15°. In that article, we find justification for introducing a positive number δ such that, for any numbers t and $\bar{t}$ in I, if $|\bar{t} - t| \leq \delta$ then $\|K_{\bar{t}} - K_t\| \leq \epsilon$, so that:

$$\begin{aligned}
\frac{1}{d}|\mathbf{K}(g)(\bar{t}) - \mathbf{K}(g)(t)| &= |\langle\!\langle\, K_{\bar{t}} - K_t, \bar{g} \,\rangle\!\rangle| \\
&\leq \langle\!\langle K_{\bar{t}} - K_t \rangle\!\rangle \langle\!\langle \bar{g} \rangle\!\rangle \\
&\leq \|K_{\bar{t}} - K_t\| \langle\!\langle \bar{g} \rangle\!\rangle \\
&\leq \epsilon \langle\!\langle g \rangle\!\rangle
\end{aligned}$$

44° At this point, let us apply relations (14) and (15), together with the celebrated Ascoli/Arzela Theorem, to derive a critical lemma. Let:

$$G: \qquad g_1, g_2, g_3, \ \ldots, \ g_j, \ \ldots$$

be a sequence in $\mathbf{C}^0$ and let:

$$F: \qquad f_1, f_2, f_3, \ \ldots, \ f_j, \ \ldots$$

be the image sequence defined by $\mathbf{K}$:

$$f_j = \mathbf{K}(g_j)$$

Let us assume that G is bounded relative to the integral norm. By relation (14), we see that F is bounded relative to the uniform norm. Moreover, by relation (15), we see that the terms of F compose a uniformly equicontinuous family. We mean to say that, for each positive number $\bar{\epsilon}$, there is a positive number δ such that, for any index j and for any numbers t and $\bar{t}$ in I, if $|\bar{t} - t| \leq \delta$ then:

$$|f_j(\bar{t}) - f_j(t)| \leq \bar{\epsilon}$$

These properties of F are precisely the stated hypotheses of the Ascoli/Arzela Theorem. By that theorem, we obtain a subsequence:

$$\Phi: \qquad \phi_1, \phi_2, \phi_3, \ \ldots, \ \phi_j, \ \ldots$$

of F such that Φ is convergent in $\mathbf{C}^0$ relative to the uniform norm.

45° Let us take a moment to sketch the proof of the Ascoli/Arzela Theorem. Let D be a countable dense subset of I. Since the sequence F is bounded relative to the uniform norm, we may apply the Bolzano/Weierstrass Theorem and the Cantor Diagonal Argument to produce a subsequence:

$$\Phi : \qquad \phi_1, \phi_2, \phi_3, \ \ldots, \ \phi_j, \ \ldots$$

of F such that, for each u in D, the sequence:

$$\phi_1(u), \phi_2(u), \phi_3(u), \ \ldots, \ \phi_j(u), \ \ldots$$

is convergent in $\mathcal{C}$. We contend that Φ is cauchy, hence convergent in $\mathbf{C}^0$ relative to the uniform norm.

46° Let $\bar{\epsilon}$ be any positive number. Since the terms of Φ compose a uniformly equicontinuous family, we may introduce a positive number δ such that, for any index j and for any numbers t and $\bar{t}$ in I, if $|\bar{t} - t| \leq \delta$ then:

$$|\phi_j(\bar{t}) - \phi_j(t)| \leq \bar{\epsilon}/3$$

Since D is dense in I, we may introduce a finite subset $D_\circ$ of D such that, for any number t in I, there is some number u in $D_\circ$ such that $|t - u| \leq \delta$. In turn, we may introduce an index n such that, for any indices p and q, if $n \leq p$ and $n \leq q$ then, for any number u in $D_\circ$, $|\phi_p(u) - \phi_q(u)| \leq \bar{\epsilon}/3$.

47° Now let p and q be any indices for which $n \leq p$ and $n \leq q$. Let t be any number in I. Let u be a number in $D_\circ$ such that $|t - u| \leq \delta$. We find that:

$$\begin{aligned} |\phi_p(t) - \phi_q(t)| &\leq |\phi_p(t) - \phi_p(u)| + |\phi_p(u) - \phi_q(u)| + |\phi_q(u) - \phi_q(t)| \\ &\leq \bar{\epsilon} \end{aligned}$$

Consequently, Φ is cauchy in $\mathbf{C}^0$ relative to the uniform norm.

48° For the analysis of $\mathbf{K}$, we require one more technical fact. Let $\mathbf{T}$ be a linear subspace of $\mathbf{C}^0$. We define the uniform norm and the integral norm for the restriction of $\mathbf{K}$ to $\mathbf{T}$:

$$\begin{aligned} \|\mathbf{K}\|_{\mathbf{T}} &\equiv sup\,\{\, \langle\!\langle \mathbf{K}(g) \rangle\!\rangle : g \in \mathbf{T}, \langle\!\langle g \rangle\!\rangle = 1 \,\} \\ \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}} &\equiv sup\,\{\, |\langle\!\langle \mathbf{K}(g), g \,\rangle\!\rangle| : g \in \mathbf{T}, \langle\!\langle g \rangle\!\rangle = 1 \,\} \end{aligned}$$

We contend that the two norms are equal:

$$(16) \qquad \|\mathbf{K}\|_{\mathbf{T}} = \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}}$$

This relation plays a basic role in the following argument.

49° To prove the contention, let us note that $\langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}} \leq \|\mathbf{K}\|_{\mathbf{T}}$ and that, for each function g in $\mathbf{C}^0$, $\langle\!\langle \mathbf{K}(g) \rangle\!\rangle \leq \|\mathbf{K}\|_{\mathbf{T}} \langle\!\langle g \rangle\!\rangle$ and $|\langle\!\langle \mathbf{K}(g), g \rangle\!\rangle| \leq \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}} \langle\!\langle g \rangle\!\rangle^2$. In turn, let g be any function in $\mathbf{T}$ for which $\langle\!\langle g \rangle\!\rangle = 1$. Let $f \equiv \mathbf{K}(g)$. Let v be any nonzero real number and let:

$$\begin{aligned} h_1 &\equiv vg + v^{-1} f \\ h_2 &\equiv vg - v^{-1} f \end{aligned}$$

We find that:

$$\begin{aligned} \langle\!\langle f \rangle\!\rangle^2 &= (1/4)\big(\langle\!\langle \mathbf{K}(h_1), h_1 \rangle\!\rangle - \langle\!\langle \mathbf{K}(h_2), h_2 \rangle\!\rangle\big) \\ &\leq (1/4)\big(\langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}} \langle\!\langle h_1 \rangle\!\rangle^2 + \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}} \langle\!\langle h_2 \rangle\!\rangle^2\big) \\ &= \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}} (1/2)(v^2 + v^{-2} \langle\!\langle f \rangle\!\rangle^2 \end{aligned}$$

Of course, the infimum for the set:

$$(1/2)(w + w^{-1} \langle\!\langle f \rangle\!\rangle^2) \qquad (0 < w)$$

of real numbers is $\langle\!\langle f \rangle\!\rangle$. It follows that $\langle\!\langle f \rangle\!\rangle \leq \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}}$. Hence, $\|\mathbf{K}\|_{\mathbf{T}} \leq \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}}$, so $\|\mathbf{K}\|_{\mathbf{T}} = \langle\!\langle \mathbf{K} \rangle\!\rangle_{\mathbf{T}}$.

50° Now we complete our analysis of $\mathbf{K}$ by induction. Let $\mathbf{T}_1 = \mathbf{C}^0$ and let $\|\mathbf{K}\|_1$ and $\langle\!\langle \mathbf{K} \rangle\!\rangle_1$ be the corresponding uniform and integral norms, respectively, for the restriction of $\mathbf{K}$ to $\mathbf{T}_1$. By relation (16), there must be a sequence:

$$G: \qquad g_1, g_2, g_3, \ \ldots, \ g_j, \ \ldots$$

in $\mathbf{T}_1$ such that, for each index j, $\langle\!\langle g_j \rangle\!\rangle = 1$ and such that the sequence:

$$|\langle\!\langle \mathbf{K}(g_1), g_1 \rangle\!\rangle|, |\langle\!\langle \mathbf{K}(g_2), g_2 \rangle\!\rangle|, |\langle\!\langle \mathbf{K}(g_3), g_3 \rangle\!\rangle|, \ \ldots \ , |\langle\!\langle \mathbf{K}(g_j), g_j \rangle\!\rangle|, \ \ldots$$

converges to $\|\mathbf{K}\|_1$. Of course, we may as well assume that the sequence:

$$\langle\!\langle \mathbf{K}(g_1), g_1 \rangle\!\rangle, \langle\!\langle \mathbf{K}(g_2), g_2 \rangle\!\rangle, \langle\!\langle \mathbf{K}(g_3), g_3 \rangle\!\rangle, \ \ldots \ , \langle\!\langle \mathbf{K}(g_j), g_j \rangle\!\rangle, \ \ldots$$

itself is convergent:

$$\lim_{j \to \infty} \langle\!\langle \mathbf{K}(g_j), g_j \rangle\!\rangle = \kappa_1$$

where $|\kappa_1| = \|\mathbf{K}\|_1$. Clearly, $0 < |\kappa_1|$. For each index j, we have:

$$\begin{aligned} 0 &\leq \langle\!\langle \mathbf{K}(g_j) - \kappa_1 g_j \rangle\!\rangle^2 \\ &= \langle\!\langle \mathbf{K}(g_j) \rangle\!\rangle^2 - 2\kappa_1 \langle\!\langle \mathbf{K}(g_j), g_j \rangle\!\rangle + \kappa_1^2 \langle\!\langle g_j \rangle\!\rangle^2 \\ &\leq \|\mathbf{K}\|_1^2 - 2\kappa_1 \langle\!\langle \mathbf{K}(g_j), g_j \rangle\!\rangle + \kappa_1^2 \\ &= 2\kappa_1 \big(\kappa_1 - \langle\!\langle \mathbf{K}(g_j), g_j \rangle\!\rangle\big) \end{aligned}$$

Consequently:

$$\lim_{j\to\infty} \langle\!\langle \mathbf{K}(g_j) - \kappa_1 g_j \rangle\!\rangle = 0 \tag{17}$$

Let:

$$F: \qquad f_1, f_2, f_3, \ \ldots, \ f_j, \ \ldots$$

be the image sequence defined from G by $\mathbf{K}$:

$$f_j = \mathbf{K}(g_j)$$

By the Ascoli/Arzela Theorem, we may introduce a subsequence:

$$\Phi: \qquad \phi_1, \phi_2, \phi_3, \ \ldots, \ \phi_j, \ \ldots$$

of F which is convergent in $\mathbf{C}^0$ relative to the uniform norm, hence, relative to the integral norm as well. Naturally, we may introduce a subsequence:

$$\Psi: \qquad \psi_1, \psi_2, \psi_3, \ \ldots, \ \psi_j, \ \ldots$$

of G such that, for each index j:

$$\mathbf{K}(\psi_j) = \phi_j$$

From relation (17), we infer that Ψ is convergent relative to the integral norm. Let h_1 be the limit of Ψ. We conclude that:

$$\mathbf{K}(h_1) = \kappa_1 h_1$$

Obviously, $\langle\!\langle h_1 \rangle\!\rangle = 1$.

51° Let $\mathbf{T}_2$ be the linear subspace of $\mathbf{C}^0$ consisting of all functions g which are orthogonal to h_1:

$$\mathbf{T}_2 = \{ \, g \in \mathbf{C}^0 \, : \langle\!\langle g, h_1 \rangle\!\rangle = 0 \, \}$$

and let $\|\mathbf{K}\|_2$ and $\langle\!\langle \mathbf{K} \rangle\!\rangle_2$ be the corresponding uniform and integral norms, respectively, for the restriction of $\mathbf{K}$ to $\mathbf{T}_2$. Imitating the foregoing argument, we obtain a number κ_2 and a function h_2 in $\mathbf{T}_2$ such that $|\kappa_2| = \|\mathbf{K}\|_2$, such that:

$$\mathbf{K}(h_2) = \kappa_2 h_2$$

and such that $\langle\!\langle h_2 \rangle\!\rangle = 1$. Obviously, $0 < |\kappa_2| \leq |\kappa_1|$.

52° Let $\mathbf{T}_3$ be the linear subspace of $\mathbf{C}^0$ consisting of all functions g which are orthogonal to h_1 and h_2:

$$\mathbf{T}_3 = \{ \, g \in \mathbf{C}^0 \, : \langle\!\langle g, h_1 \rangle\!\rangle = 0, \, \langle\!\langle g, h_2 \rangle\!\rangle = 0 \, \}$$

and let $\|\mathbf{K}\|_3$ and $\langle\!\langle \mathbf{K} \rangle\!\rangle_3$ be the corresponding uniform and integral norms, respectively, for the restriction of $\mathbf{K}$ to $\mathbf{T}_3$. Imitating the foregoing argument, we obtain a number κ_3 and a function h_3 in $\mathbf{T}_3$ such that $|\kappa_3| = \|\mathbf{K}\|_3$, such that:

$$\mathbf{K}(h_3) = \kappa_3 h_3$$

and such that $\langle\!\langle h_3 \rangle\!\rangle = 1$. Obviously, $0 < |\kappa_3| \leq |\kappa_2| \leq |\kappa_1|$.

53° Continuing inductively, we obtain a sequence:

$$K : \qquad \kappa_1, \kappa_2, \kappa_3, \ \dots \ , \kappa_j, \ \dots$$

of (nonzero) real numbers and a sequence:

$$H : \qquad h_1, h_2, h_3, \ \dots \ , h_j, \ \dots$$

of functions in $\mathbf{C}^0$ satisfying the relations just described.

54° We contend that:

$$|\kappa_j| \downarrow 0 \tag{18}$$

Let us suppose to the contrary that there is a positive number ϵ such that, for each index j:

$$\epsilon \leq |\kappa_j|$$

It would follow that:

$$\langle\!\langle \kappa_j^{-1} h_j \rangle\!\rangle \leq \epsilon^{-1} \quad \text{and} \quad \mathbf{K}(\kappa_j^{-1} h_j) = h_j$$

By the Ascoli/Arzela Theorem, we may introduce a subsequence of H which is convergent in $\mathbf{C}^0$ relative to the uniform norm, hence, relative to the integral norm as well. However, for any indices j and k, if $j \neq k$ then $\langle\!\langle h_j, h_k \rangle\!\rangle = 0$, so:

$$\langle\!\langle h_j - h_k \rangle\!\rangle^2 = 2$$

By this contradiction, we infer that our supposition is false. Hence:

$$|\kappa_j| \downarrow 0$$

55° Let us replace the sequence K by the sequence:

$$\Lambda : \qquad \lambda_1, \lambda_2, \lambda_3, \ \dots \ , \lambda_j, \ \dots$$

of real numbers defined by inversion. That is, for each index j, $\lambda_j = \kappa_j^{-1}$. Of course:

$$\mathbf{L}(h_j) = \lambda_j h_j$$

56° Now we have succeeded in producing sequences Λ and H which meet conditions $(B1)$, $(B2)$, and $(B3)$ in Theorem B. Let us prove that condition $(B4)$ holds as well.

57° Let f be any function in $\mathbf{S}$. Let $g \equiv \mathbf{L}(f)$ so $\mathbf{K}(g) = f$. For each index k, let g_k be the function in $\mathbf{C}^0$ defined as follows:

$$g_k \equiv g - \sum_{j=1}^{k} \langle\!\langle\, g, h_j \,\rangle\!\rangle h_j$$

Let G be the sequence in $\mathbf{C}^0$ so defined:

$$G : \qquad g_1, g_2, g_3, \ \ldots, \ g_k, \ \ldots$$

Let F be the image sequence:

$$F : \qquad f_1, f_2, f_3, \ \ldots, \ f_k, \ \ldots$$

defined by $\mathbf{K}$:

$$f_k \equiv \mathbf{K}(g_k)$$

We contend that F converges in $\mathbf{C}^0$ to 0, relative to the uniform norm. If that were so, then the series:

$$\sum_{j=1}^{\infty} \langle\!\langle\, f, h_j \,\rangle\!\rangle h_j$$

would converge in $\mathbf{C}^0$ to f relative to the uniform norm, because:

$$\begin{aligned} f - \sum_{j=1}^{k} \langle\!\langle\, f, h_j \,\rangle\!\rangle h_j &= \mathbf{K}(g) - \sum_{j=1}^{k} \langle\!\langle\, \mathbf{K}(g), h_j \,\rangle\!\rangle h_j \\ &= \mathbf{K}(g) - \sum_{j=1}^{k} \langle\!\langle\, g, \mathbf{K}(h_j) \,\rangle\!\rangle h_j \\ &\qquad\qquad\qquad\qquad \text{(since } \mathbf{K}(h_j) = \kappa_j h_j) \\ &= \mathbf{K}(g) - \sum_{j=1}^{k} \langle\!\langle\, g, h_j \,\rangle\!\rangle \mathbf{K}(h_j) \\ &= \mathbf{K}(g_k) \\ &= f_k \end{aligned}$$

These observations form a precise statement of condition $(B4)$.

58° Let us prove the foregoing contention. Obviously, for each index j, if $1 \leq j \leq k$ then $\langle\!\langle\, g_k, h_j \,\rangle\!\rangle = 0$, so:

$$\langle\!\langle g_k \rangle\!\rangle^2 = \langle\!\langle g \rangle\!\rangle^2 - \sum_{j=1}^{k} |\langle\!\langle\, g, g_j \,\rangle\!\rangle|^2 \leq \langle\!\langle g \rangle\!\rangle^2$$

In particular, g_k must lie in $\mathbf{T}_{k+1}$ and:

$$\langle\!\langle \mathbf{K}(g_k) \rangle\!\rangle \leq \|\mathbf{K}\|_{k+1} \langle\!\langle g_k \rangle\!\rangle = |\kappa_{k+1}| \langle\!\langle g \rangle\!\rangle$$

By relation (18):

$$\lim_{k \to \infty} \langle\!\langle \mathbf{K}(g_k) \rangle\!\rangle = 0 \qquad (19)$$

Consequently, F converges in $\mathbf{C}^0$ to 0, relative to the integral norm, but that is not enough.

59° We must prove the stronger assertion, namely, that F converges in $\mathbf{C}^0$ to 0, relative to the uniform norm. To that end, we invoke a Peculiar but sometimes useful Principle from the theory of convergence in metric spaces. It states that:

> *For every sequence and for every point, the sequence converges to the point iff, for every subsequence of the given sequence, there is a subsequence of the subsequence which converges to the point.*

Arguing by contradiction, one can prove the principle very easily.

60° Let us apply the principle. By the Ascoli/Arzela Theorem, every subsequence of F must itself admit a subsequence which is convergent in $\mathbf{C}^0$ relative to the uniform norm. By relation (19), the limit of that *sub-subsequence* must be 0. By the Peculiar Principle, F must converge in $\mathbf{C}^0$ to 0, relative to the uniform norm. •

The Sturm/Liouville Problem

61° From Theorem B, we see that the analysis of the Symmetric operator $\mathbf{L}'$ reduces to two steps:

(•) find the eigenvalues λ for $\mathbf{L}'$

(•) for each such λ, find an orthonormal basis of eigenfunctions for the (at most two dimensional) eigenspace $\mathbf{\Lambda}$ corresponding to λ

62° One should arrange the eigenvalues and eigenfunctions in a reasonable order, side by side. For example, one might list the eigenvalues in order of increasing absolute value. When two distinct eigenvalues have the same absolute value, one might list the negative value first. For the cases (if any) in which the corresponding eigenspace is two dimensional, one would list the eigenvalue twice, in order to make room for two corresponding eigenfunctions, which form a basis for the eigenspace.

63° One might wonder whether the list Λ of eigenvalues displayed in Theorem B is complete. Indeed, let λ be a real number and let f be a function in $\mathbf{S}$ such that $f \neq 0$ and such that $\mathbf{L}(f) = \lambda f$. Of course, $\lambda \neq 0$. By $(B4)$, there must be some index j such that $\langle\!\langle\, f, h_j \,\rangle\!\rangle \neq 0$. Let k be such an index. By simple steps, we find that:

$$\lambda \langle\!\langle\, f, h_k \,\rangle\!\rangle = \lambda_k \langle\!\langle\, f, h_k \,\rangle\!\rangle$$

Hence, $\lambda = \lambda_k$.

Fourier Series

64° Let us continue consideration of the illustration described in articles 30° and 34°. In this case, we may display the eigenfunctions as a bilateral array:

$$h_k(t) = exp(ikt) \qquad (k \in \mathcal{Z})$$

For each k in $\mathcal{Z}$, the corresponding eigenvalue is:

$$\lambda = 1 + k^2$$

65° Theorem B yields the celebrated Theorem of Fourier. For each complex valued function f defined on the interval $[\,-\pi, \pi\,]$, if f is twice continuously differentiable, if $f(-\pi) = f(\pi)$, and if $f^\circ(-\pi) = f^\circ(\pi)$ then f is the uniform limit of its Fourier Series:

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \, exp(ikt)$$

with Fourier Coefficients:

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, exp(-ikt) dt$$

Chapter 4 FRIEDMANN'S EQUATION

Introduction

1° In this chapter, we show that the Equations of Newton governing Celestial Mechanics admit solutions which serve as models for an Expanding Universe of material particles. We find that, for certain initial configurations, the particles may expand and contract in concert. Remarkably, the corresponding time dependent scale factor must satisfy the celebrated Equation of Friedmann, the same equation which figures in Modern Cosmology.

The Equations of Newton

2° Let n be a positive integer $(3 \leq n)$ and let j and k be generic positive integers for which $1 \leq j \leq n$ and $1 \leq k \leq n$. Let us introduce the following notation:

$$\begin{aligned} 0 &< m_j \\ \mathbf{r}_j &= (r_j^1, r_j^2, r_j^3) \\ \mathbf{r} &= (\mathbf{r}_1, \mathbf{r}_2, \ \ldots \ , \mathbf{r}_n) \\ \rho_{jk}(\mathbf{r}) &= \|\mathbf{r}_j - \mathbf{r}_k\| \\ \rho(\mathbf{r}) &= \min_{j \neq k} \rho_{jk}(\mathbf{r}) \end{aligned}$$

Let Λ be any number in $\mathcal{R}$. Let us introduce the following Autonomous Second Order ODE, in the $3n$ variables $\mathbf{r}$:

$$(\mathbf{N}) \qquad \begin{aligned} m_1 \mathbf{r}_1^{\circ\circ} &= \sum_{j \neq 1} \frac{m_j m_1}{\rho_{j1}(\mathbf{r})^2} \frac{\mathbf{r}_j - \mathbf{r}_1}{\rho_{j1}(\mathbf{r})} + \frac{1}{3}\Lambda m_1 \mathbf{r}_1 \\ m_2 \mathbf{r}_2^{\circ\circ} &= \sum_{j \neq 2} \frac{m_j m_2}{\rho_{j2}(\mathbf{r})^2} \frac{\mathbf{r}_j - \mathbf{r}_2}{\rho_{j2}(\mathbf{r})} + \frac{1}{3}\Lambda m_2 \mathbf{r}_2 \\ &\vdots \\ m_n \mathbf{r}_n^{\circ\circ} &= \sum_{j \neq n} \frac{m_j m_n}{\rho_{jn}(\mathbf{r})^2} \frac{\mathbf{r}_j - \mathbf{r}_n}{\rho_{jn}(\mathbf{r})} + \frac{1}{3}\Lambda m_n \mathbf{r}_n \end{aligned} \qquad (0 < \rho(\mathbf{r}))$$

These are Newton's Equations, in natural units of measurement, governing a system of n particles in $\mathcal{R}^3$ having masses $m_1, m_2, \ \ldots \ , m_n$.

3° The particles interact as usual by mutual gravitational attraction but suffer in addition a *cosmological force* defined by Λ.

4° We imagine that the particles are *stars*, even *galaxies*.

5° Let:

$$\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2, \ \ldots \ , \mathbf{s}_n)$$

be any vector in $\mathcal{R}^{3n}$ for which $0 < \rho(\mathbf{s})$. Let J be any open interval in $\mathcal{R}$ such that $0 \in J$ and let S be a function defined on J with values in $\mathcal{R}^+$. Let σ be any number in $\mathcal{R}$. We assume that:

$$(I) \qquad S(0) = 1, \qquad S^\circ(0) = \sigma$$

Let $\mathbf{r}$ be the function with values in $\mathcal{R}^{3n}$, defined on J as follows:

$$(*) \qquad \mathbf{r}(t) = S(t)\mathbf{s} \qquad (t \in J)$$

We inquire whether $\mathbf{r}$ may satisfy Newton's Equations $(\mathbf{N})$. We will find that it can be so, provided that the initial configuration $\mathbf{s}$ meets a certain stringent condition, namely, that it be a Central Configuration, and provided that the scale factor S meets a certain celebrated Autonomous Second Order ODE, namely, the Equation of Friedmann.

Central Configurations

6° For the function $\mathbf{r}$ defined by relation $(*)$, we may recast equations $(\mathbf{N})$ as follows:

$$(S) \qquad \begin{aligned} S^{\circ\circ} m_1 \mathbf{s}_1 &= S^{-2} \sum_{j \neq 1} \frac{m_j m_1}{\rho_{j1}(\mathbf{s})^2} \frac{\mathbf{s}_j - \mathbf{s}_1}{\rho_{j1}(\mathbf{s})} + S\frac{1}{3}\Lambda m_1 \mathbf{s}_1 \\ S^{\circ\circ} m_2 \mathbf{s}_2 &= S^{-2} \sum_{j \neq 2} \frac{m_j m_2}{\rho_{j2}(\mathbf{s})^2} \frac{\mathbf{s}_j - \mathbf{s}_2}{\rho_{j2}(\mathbf{s})} + S\frac{1}{3}\Lambda m_2 \mathbf{s}_2 \\ &\vdots \\ S^{\circ\circ} m_n \mathbf{s}_n &= S^{-2} \sum_{j \neq n} \frac{m_j m_n}{\rho_{jn}(\mathbf{s})^2} \frac{\mathbf{s}_j - \mathbf{s}_n}{\rho_{jn}(\mathbf{s})} + S\frac{1}{3}\Lambda m_n \mathbf{s}_n \end{aligned}$$

Let us evaluate the foregoing equations (S) at $t = 0$. We find that there must exist a certain number $\tau(\mathbf{s})$, namely:

$$\tau(\mathbf{s}) = S^2(0)S^{\circ\circ}(0) - \frac{1}{3}\Lambda S^3(0) = S^{\circ\circ}(0) - \frac{1}{3}\Lambda$$

such that:

$$(C) \qquad \begin{aligned} \sum_{j \neq 1} \frac{m_j m_1}{\rho_{j1}(\mathbf{s})^2} \frac{\mathbf{s}_j - \mathbf{s}_1}{\rho_{j1}(\mathbf{s})} &= \tau(\mathbf{s}) m_1 \mathbf{s}_1 \\ \sum_{j \neq 2} \frac{m_j m_2}{\rho_{j2}(\mathbf{s})^2} \frac{\mathbf{s}_j - \mathbf{s}_2}{\rho_{j2}(\mathbf{s})} &= \tau(\mathbf{s}) m_2 \mathbf{s}_2 \\ &\vdots \\ \sum_{j \neq n} \frac{m_j m_n}{\rho_{jn}(\mathbf{s})^2} \frac{\mathbf{s}_j - \mathbf{s}_n}{\rho_{jn}(\mathbf{s})} &= \tau(\mathbf{s}) m_n \mathbf{s}_n \end{aligned}$$

Under this condition, one says that $\mathbf{s}$ is a *central configuration*, with *scale factor* $\tau(\mathbf{s})$. It turns out that $\tau(\mathbf{s})$ must be negative. See article 18°.

7° Adding equations (C), we find that:

$$\frac{1}{\mu}\sum_j m_j \mathbf{r}_j = \mathbf{0} \qquad (\mu = \sum_j m_j)$$

Friedmann's Equation

8° Granted equations (C), we may recast equations (S) as an Autonomous Second Order ODE:

$$(\mathbf{F}) \qquad S^{\circ\circ} = \tau S^{-2} + \frac{1}{3}\Lambda S \qquad (0 < S)$$

where $\tau = \tau(\mathbf{s})$. Remarkably, equation $(\mathbf{F})$ coincides with the Equation of Friedmann, derived in modern studies of Cosmology. It also coincides with Newton's Equation for radial motion, modified by the cosmological force defined by Λ.

9° With reference to article 5°, we conclude that the function:

$$(*) \qquad \mathbf{r}(t) = S(t)\mathbf{s} \qquad (t \in J)$$

satisfies Newton's Equations $(\mathbf{N})$ iff $\mathbf{s}$ is a central configuration, with scale factor $\tau(\mathbf{s})$, and S satisfies Friedmann's Equation $(\mathbf{F})$, where $\tau = \tau(\mathbf{s})$.

10° Let us note that, for each j $(1 \leq j \leq n)$:

$$\mathbf{r}^\circ_j(t) = S^\circ(t)\mathbf{s}_j = \frac{S^\circ(t)}{S(t)} S(t)\mathbf{s}_j = \frac{S^\circ(t)}{S(t)}\mathbf{r}_j(t)$$

Hence:

$$\mathbf{r}^\circ_j(t) = H(t)\mathbf{r}_j(t)$$

Consequently, the velocities of the particles are proportional to their positions. One may refer to the common factor of proportionality as the Hubble Parameter:

$$H(t) = \frac{S^\circ(t)}{S(t)}$$

11° The ODE (**F**) is equivalent to the following Autonomous First Order ODE:

$$(\mathbf{G}) \qquad \begin{aligned} S^\circ &= T \\ T^\circ &= \tau S^{-2} + \frac{1}{3}\Lambda S \end{aligned}$$

12° Let h be the function defined as follows:

$$h(S,T) = \frac{1}{2}T^2 + \tau S^{-1} - \frac{1}{6}\Lambda S^2 \qquad (0 < S,\ T \in \mathcal{R})$$

We find that:

$$\nabla h = (-\tau S^{-2} - \frac{1}{3}\Lambda S,\ T)$$

Obviously, the vector field which defines (**G**) and the gradient field ∇h are orthogonal:

$$(T,\ \tau S^{-2} + \frac{1}{3}\Lambda S) \bullet (-\tau S^{-2} - \frac{1}{3}\Lambda S,\ T) = 0$$

Let γ be a maximum integral curve for (**G**):

$$\gamma(t) = (S(t), T(t)) \qquad (t \in J)$$

By the foregoing observation, the function:

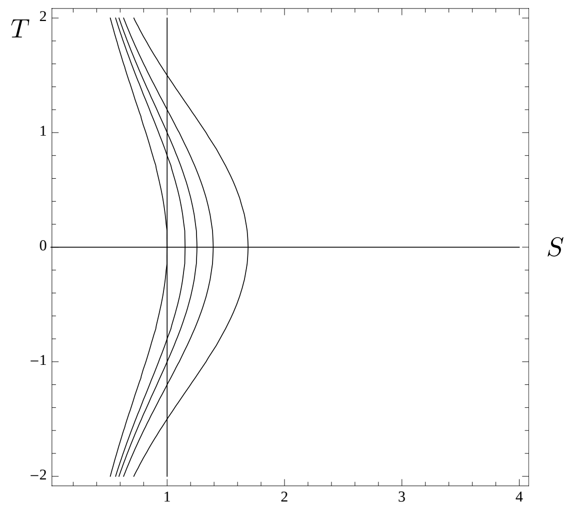
$$h(S(t), T(t)) \qquad (t \in J)$$

must be constant. Hence, $\gamma(J)$ must be a subset of one of the level sets for h.

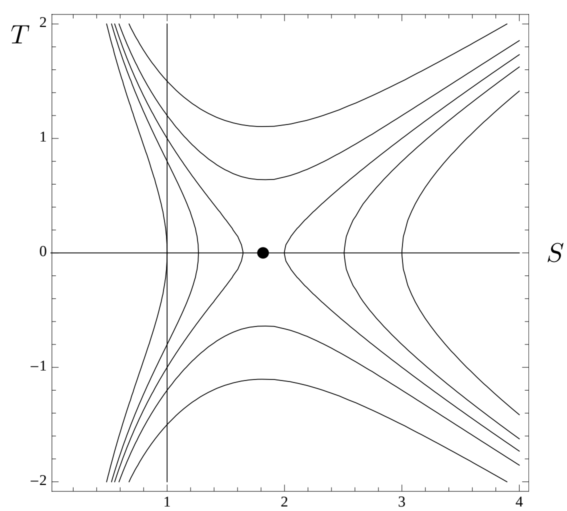
13° For the cases in which $0 < \Lambda$, we find exactly one critical point for (**G**):

$$((-\frac{3\tau}{\Lambda})^{1/3},\ 0)$$

14° By studying the following diagrams of level sets for h, one may visualize some of the solutions of Friedmann's Equation. In particular, one may see scenarios involving "contraction, expansion, capture, and escape."



$\Lambda = -1, \quad \tau = -2$



$\Lambda = 1, \quad \tau = -2$

Lagrange Multipliers

15° Let us show that central configurations exist. To that end, let us introduce the (negative of the) potential energy function for the system of particles:

$$U(\mathbf{r}) = \sum_{j<k}\sum \frac{m_j m_k}{\rho_{jk}(\mathbf{r})} \qquad (0 < \rho(\mathbf{r}))$$

One can easily verify that:

$$
\begin{aligned}
(\nabla_1 U)(\mathbf{r}) &= \sum_{j \neq 1} \frac{m_j m_1}{\rho_{j1}(\mathbf{r})^2} \frac{\mathbf{r}_j - \mathbf{r}_1}{\rho_{j1}(\mathbf{r})} \\
(\nabla_2 U)(\mathbf{r}) &= \sum_{j \neq 2} \frac{m_j m_2}{\rho_{j2}(\mathbf{r})^2} \frac{\mathbf{r}_j - \mathbf{r}_2}{\rho_{j2}(\mathbf{r})} \\
&\vdots \\
(\nabla_n U)(\mathbf{r}) &= \sum_{j \neq n} \frac{m_j m_n}{\rho_{jn}(\mathbf{r})^2} \frac{\mathbf{r}_j - \mathbf{r}_n}{\rho_{jn}(\mathbf{r})}
\end{aligned}
$$

In turn, let us introduce (one half) the moment of inertia function for the system of particles:

$$
\Omega(\mathbf{r}) = \frac{1}{2} \sum_j m_j \|\mathbf{r}_j\|^2
$$

Obviously:

$$
\begin{aligned}
(\nabla_1 \Omega)(\mathbf{r}) &= m_1 \mathbf{r}_1 \\
(\nabla_2 \Omega)(\mathbf{r}) &= m_2 \mathbf{r}_2 \\
&\vdots \\
(\nabla_n \Omega)(\mathbf{r}) &= m_n \mathbf{r}_n
\end{aligned}
$$

In the foregoing relations, ∇_k signals the k-th gradient operator:

$$
\nabla_k = (\frac{\partial}{\partial r_k^1}, \frac{\partial}{\partial r_k^2}, \frac{\partial}{\partial r_k^3})
$$

16° Let us introduce an arbitrary vector $\mathbf{r}^*$ in $\mathcal{R}^{3n}$ for which $0 < \rho(\mathbf{r}^*)$. Let W be the corresponding (modified) level set for Ω:

$$
W: \qquad \Omega(\mathbf{r}) = \Omega(\mathbf{r}^*) \qquad (0 < \rho(\mathbf{r}))
$$

We contend that there exists a vector $\mathbf{s}$ in W at which U achieves its minimum value on W. We prove the contention as follows. Let $u = U(\mathbf{r}^*)$. Of course, $0 < u$. For any indices j and k $(1 \leq j < k \leq n)$, let ϵ_{jk} be a positive number such that, for any vector $\mathbf{r}$ in W, if $\rho_{jk}(\mathbf{r}) < \epsilon_{jk}$ then $u < U(\mathbf{r})$. Let $\bar{W}$ be the subset of W consisting of all vectors $\mathbf{r}$ such that, for any indices j and k $(1 \leq j < k \leq n)$, $\epsilon_{jk} \leq \rho_{jk}(\mathbf{r})$. Of course, $\bar{W}$ is compact. Let $\mathbf{s}$ be a vector in $\bar{W}$ at which U achieves its minimum value on $\bar{W}$. By design, $U(\mathbf{s})$ is the minimum value of U on W.

17° Now we may apply the theory of Lagrange Multipliers to introduce a number $\tau(\mathbf{s})$ in $\mathcal{R}$ such that the following equations are satisfied:

$$(L) \qquad \begin{aligned} (\nabla_1 U)(\mathbf{s}) &= \tau(\mathbf{s})(\nabla_1 \Omega)(\mathbf{s}) \\ (\nabla_2 U)(\mathbf{s}) &= \tau(\mathbf{s})(\nabla_2 \Omega)(\mathbf{s}) \\ &\vdots \\ (\nabla_n U)(\mathbf{s}) &= \tau(\mathbf{s})(\nabla_n \Omega)(\mathbf{s}) \end{aligned}$$

Consequently, $\mathbf{s}$ is a central configuration.

18° Let us note that U is homogeneous of degree -1. We may apply Euler's Theorem to obtain:

$$\sum_j (\nabla_j U)(\mathbf{s}) \bullet \mathbf{s}_j = -U(\mathbf{s})$$

Obviously:

$$\sum_j (\nabla_j \Omega)(\mathbf{s}) \bullet \mathbf{s}_j = 2\,\Omega(\mathbf{s})$$

Now, by equation (L), we find that:

$$2\tau(\mathbf{s})\Omega(\mathbf{s}) = -U(\mathbf{s})$$

Hence:

$$\tau(\mathbf{s}) = -\frac{1}{2}\frac{U(\mathbf{s})}{\Omega(\mathbf{s})} < 0$$