Studies of Non-homogeneous Cosmological Models

Thesis by

Guy C. Omer, Jr.

In Partial Fulfilment of the Requirements

for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1947

Acknowledgment

The author would like to express his appreciation to Prof. Richard C. Tolman for his continued help and encouragement. It was Dr. Tolman who suggested this problem and who has followed the evolution of the material which is presented here.

Abstract

A family of spherically symmetric, non-static, non-homogeneous relativistic models with zero pressure is considered. In the first chapter the pioneer work of Lemaitre and Tolman is reviewed. The fundamental partial differential equation is transformed into a parametric pair which may be solved explicitly in terms of known functions.

In the second chapter the nature of the most general solutions and the conditions under which they will exist are determined. The Robertson-Tolman notation is extended to apply to non-homogeneous models. The fundamental equation is solved in terms of Weierstrassian elliptic functions. The solutions in terms of elliptic functions are then shown to behave in the expected manner.

Approximate solutions not involving elliptic functions which will apply when the cosmological constant is small as compared to another parameter are derived in the following chapter. Further information about the relations between the various solutions is obtained.

In the fourth chapter all of the special non-elliptic solutions are found for which the coefficients of the fundamental equation are finite. The well known static and non-static homogeneous models with zero pressure are shown to be special cases of this family of cosmological

models. The Robertson two-dimensional graphical representation of the existence conditions for homogeneous models is extended to its equivalent three-dimensional representation for these models. A general expression for the local proper density within these models is derived and further physical restrictions upon the solutions are developed.

In the final chapter the usefulness of this family of models is illustrated by applying the special solutions for a zero cosmological constant to the cosmological problem. Suitable expressions are derived for the red-shift, for the number of nebulae which would be counted to a limiting magnitude, and for the observed magnitude of a source of known luminosity located at a stated coordinate. A model is then constructed which agrees with Hubble's observational data to a first approximation.

Contents

Chapter I

Introduction

We shall consider a group of spherically symmetrical, non-homogeneous, non-static cosmological models composed entirely of dust particles which, by definition, are incapable of exerting any pressures whatsoever. The use of cosmological models filled with only dust particles is a good approximation to the known observational facts of the present epoch. In these models each dust particle represents a single nebula. No significant intergalactic forces other than gravitation are known at present. Any radiation pressures upon the nebulae within our observational neighborhood are vanishingly small. No collisions between the nebulae, which would produce a kinetic "gas" pressure, have been observed or assumed. Hence dust-filled models are a good first approximation. The energy-momentum tensor for such models will be

$$T^{\mu\nu} = \rho \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} \tag{1.1}$$

where ρ is macroscopic density of the dust as measured by a local observer moving with the dust and the quantities $\frac{dx^{\mu}}{ds}$ and $\frac{dx^{\nu}}{ds}$ are components of the velocity of the dust with respect to the coordinates in use.

In this work a special coordinate system will be used which moves with the dust. Such coordinate systems

are called "co-moving" coordinates. By definition the spacial components of the velocity of the dust with respect to co-moving coordinates shall be zero. That is, $\frac{dx^\mu}{ds} = 0$ for $\mu \neq 4$. The line-element to be used, which is yet to be developed, is one having a "cosmic" time orthogonal to the space-like components. Hence the time-like component of the velocity of the dust with respect to the co-moving coordinates shall be unity. That is, $\frac{dx^4}{ds} = 1$. Consequently the energy-momentum tensor of (1.1) reduces to

$$T^{\mu\nu} = 0 \quad [\mu \text{ or } \nu \neq 4] \qquad\qquad T^{44} = \rho \tag{1.2}$$

To put this energy-momentum tensor (1.2) into a form such that Dingle's Formulae (Ref. 1, pp. 254-7) may be used, we must lower one index. Lowering the index, we find $T^{\mu}_{\nu} = g_{\nu\alpha}T^{\alpha\mu} = 0$ for all $\mu \neq 4$ since all of the $T^{\alpha\mu}$ in the summation will be zero from (1.2). Also, $T^{4}_{\nu} = g_{\nu\alpha}T^{\alpha 4} = 0$ for all $\nu \neq 4$ since all of the $T^{\alpha 4}$ when $\alpha \neq 4$ in the summation will be zero from (1.2) while $g_{\nu 4} = 0$ from the assumed form of the line-element since it is assumed that the "cosmic" time will be orthogonal to all of the space-like components. Finally, $T^{4}_{4} = g_{4\alpha}T^{\alpha 4} = \rho$ since all of the $T^{\alpha 4}$ where $\alpha \neq 4$ in the summation will be zero from (1.2) while $g_{44} = 1$ from the assumed

form of the line element and $T^{44} = \rho$ from (1.2). Therefore the energy-momentum tensor of (1.2) with one lowered index becomes

$$T^{\mu}_{\nu} = 0 \quad [\mu \text{ or } \nu \neq 4] \qquad\qquad T^{4}_{4} = \rho \tag{1.3}$$

The energy-momentum tensor of (1.2) and (1.3) contains three assumptions. The first and most important assumption is the total absence of pressure in the universe. This is one of the basic postulates. The two remaining assumptions were made for mathematical convenience. They are the use of co-moving coordinates and the use of a line-element with an orthogonal "cosmic" time.

Since the zero pressure postulate is the important one, it might be well to examine it more closely. A pressure would be measured by a local observer moving with the dust if there were a transfer of momenta across any of the faces of a three-dimensional spacial cell that he could erect about himself for purposes of measurement. Such a transfer of momenta could be produced by two mechanisms, either by a flow of radiation, or by the motion of a material particle across a cell face. All radiation must be excluded from these models. Since radiation will move with the limiting velocity with respect to any coordinates that the local observer might use, momenta would be transferred across the cell faces. The prohibition on transfer of momenta by material particles is a more

delicate restriction. It is mathematically possible to have co-moving coordinates in which the particles "mix" upon motion. That is, the particles in the neighborhood of co-moving coordinate r_1 would "flow through" the particles in the neighborhood of co-moving coordinate r_2 during motion. A local observer at r_1 would then observe a macroscopic pressure because of the transfer of momenta by the particles in the neighborhood of r_2. Consequently, a restriction must be placed upon the types of possible motion within the models to be considered. The space within a model can be divided into unequal cells produced by connecting together adjacent particles. Then each cell, as viewed by a local observer, is free to change in shape and size, but no cell is allowed to penetrate another. This restriction, it will be seen, places certain limits on otherwise arbitrary functions of r that will arise in the solutions of the Field Equations for the models under consideration.

That "mixing" motions are gravitationally possible is illustrated by the present motion of the Solar System through the Ursa Major Group (stars of Ursa Major, Sirius, α Coronae, β Aurigae, et al). This is the type of motion on the stellar scale that must be forbidden in the models under discussion on the galactic scale.

Only one of the postulates, that of zero pressure, has been utilized so far. The other postulate, that of

spherical symmetry, will be used in establishing the line-element. In reducing the line-element, the mathematically simplifying assumptions of co-moving coordinates and of an orthogonal "cosmic" time must not be violated. The treatment used in producing the line-element will be identical to that given by Tolman (Ref. 1, pp. 364-6).

We shall start with the most general form of line-element which exhibits spacial spherical symmetry

$$ds^2 = -e^{\lambda}dr^2 - e^{\mu}\left(r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right) + e^{\nu}dt^2 + 2a\,dr\,dt \qquad (1.4)$$

where λ, μ, ν, and a are functions of r and t. We shall simplify this line-element by allowable coordinate transformations which do not upset the co-moving character of the coordinates to a form having $g_{44} = 1$ and all other $g_{\alpha 4} = 0$.

Possible coordinate transformations are restricted by the requirement that

$$\frac{dr}{ds} = \frac{d\theta}{ds} = \frac{d\phi}{ds} = 0 \qquad (1.5)$$

These are the conditions which must be satisfied by the spacial components of the particle velocities in co-moving coordinates. This means that no transformations of r, θ, or ϕ are possible. However, transformations may be made in the time-like coordinate t.

Let us transform the time-like coordinate t to a new time-like coordinate t' which will be a function of r and t. This transformation may be differentially defined by the equation

$$dt' = \eta(a\,dr + e^{\nu}dt) \tag{1.6}$$

where η is an integrating factor, a function of r and t, which makes the right-hand side of (1.6) a perfect differential. In accordance with (1.6) we shall have

$$e^{\nu}dt^2 + 2a\,dr\,dt = \frac{dt'^2}{\eta^2 e^{\nu}} - \frac{a^2}{e^{\nu}}dr^2 \tag{1.7}$$

Substituting (1.7) into (1.4) and dropping the primes, we find that the line-element is reduced to the simpler form

$$ds^2 = -e^{\lambda}dr^2 - e^{\mu}(r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2) + e^{\nu}dt^2 \tag{1.8}$$

where λ, μ, and ν are now functions of r and the present t. The co-moving relations (1.5) are not disturbed since r, θ, and ϕ are the same variables as before. The line-element (1.8) has now been reduced to one of Dingle's general forms (Ref. 1, p. 253).

We have reduced the single cross-term g_{14} to zero by the transformation (1.6) but we have not yet reduced g_{44} to unity. This further reduction can be made by a more comprehensive consideration of the co-moving character of the

coordinates. In co-moving coordinates a given particle must be stationary in its spacial coordinates. That is, the r, Θ, and ϕ coordinates for a given particle must remain constant for all past, present, and future times. It is not enough for the spacial components of the particle velocity to reduce to zero. It is also necessary for the spacial components of the particle acceleration to reduce to zero. A particle might be accelerated by many different forces, such as a pressure gradient, electro-magnetic or gravitational forces. We have not postulated any electro-magnetic forces in the models that we are considering. Since we have postulated zero pressure everywhere in our models, there will be no pressure gradients. Hence the only forces that might accelerate the particles with respect to their coordinate frames will be gravitational. We must, therefore, impose the further condition that the spacial components of the gravitational accelerations must be zero in our co-moving coordinates. This condition may be stated in terms of the equations for a geodesic, by recalling the co-moving relations of (1.5) as

$$\frac{d^2r}{ds^2} = -\left\{44,1\right\}\left(\frac{dt}{ds}\right)^2 = 0 \qquad \frac{d^2\Theta}{ds^2} = -\left\{44,2\right\}\left(\frac{dt}{ds}\right)^2 = 0 \qquad \frac{d^2\phi}{ds^2} = -\left\{44,3\right\}\left(\frac{dt}{ds}\right)^2 = 0 \tag{1.9}$$

The conditions of (1.9), it might be emphasized, are applicable only because of the absence of pressure gradients in these models.

The Christoffel three-index symbols of (1.9) are easily evaluated by the use of Dingle's Formulae (Ref. 1, p. 254) by taking $D = e^{\nu}$. By these means the conditions (1.9) will be seen to reduce readily to the conditions

$$\frac{\partial \nu}{\partial r} = \frac{\partial \nu}{\partial \theta} = \frac{\partial \nu}{\partial \phi} = 0 \tag{1.10}$$

The second and third of the conditions (1.10) are identically true since ν in the line-element (1.8) was taken to be a a function of r and t only. The first condition, however, shows that the quantity ν is a function of t only. Therefore, we may make a simple transformation in the time-scale only defined differentially as

$$dt' = e^{\frac{1}{2}\nu}\, dt \tag{1.11}$$

Since this scale transformation does not involve any of the spacial coordinates, the co-moving character of the coordinates is untouched.

By substituting the scale transformation of (1.11) into the line-element (1.8) and dropping the primes, the line-element becomes

$$ds^2 = -e^{\lambda} dr^2 - e^{\mu}\left(r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right) + dt^2 \tag{1.12}$$

where λ and μ continue as functions of r and the present t. We have now achieved our original objective

of reducing g_{44} to unity. We can, however, make one further simple reduction. Let us define a new quantity

$$e^{\omega} = e^{\mu + 2\log r} \tag{1.13}$$

by collecting terms. Substituting this new quantity (1.13) into the line-element (1.12) we have

$$ds^2 = -e^{\lambda}dr^2 - e^{\omega}(d\theta^2 + \sin^2\theta\, d\phi^2) + dt^2 \tag{1.14}$$

where λ and ω are functions of r and t. This is the non-homogeneous line-element of Tolman (Ref. 2) and Lemaitre (Ref. 3).

The line-element (1.14) contains in the simplest possible form the second postulate of spherical symmetry along with the mathematically simplifying assumptions of co-moving coordinates and an orthogonal "cosmic" time. The energy-momentum tensor (1.3) contains the first postulate of zero pressure along with the mathematically simplifying assumptions of co-moving coordinates and an orthogonal "cosmic" time. The line-element and the energy-momentum tensor are, therefore, compatible and must be combined to produce the desired series of cosmological models. The connection between the two is, of course, Einstein's field equations of general relativity

$$-8\pi T^{\mu}_{\nu} = R^{\mu}_{\nu} - \tfrac{1}{2} R g^{\mu}_{\nu} + \Lambda g^{\mu}_{\nu} \tag{1.15}$$

which equates the tensor g_ν^μ derived from the line-element with the energy-momentum tensor. Relativistic units of length, time, and mass are used in (1.15). The constant Λ in the last term is the hypothetical cosmological constant.

The equations (1.15) can be produced without great difficulty by the use of Dingle's formulae (Ref. 1, pp. 254-7) by taking

$$A = e^{\lambda} \qquad x^1 = r$$
$$B = e^{\omega} \qquad x^2 = \theta$$
$$C = e^{\omega}\sin^2\theta \qquad x^3 = \phi$$
$$D = 1 \qquad x^4 = t$$

Then

$$8\pi T_1^1 = e^{-\omega} - e^{-\lambda}\frac{\omega'^2}{4} + \ddot{\omega} + \frac{3}{4}\dot{\omega}^2 - \Lambda = 0 \tag{1.16}$$

$$8\pi T_2^2 = 8\pi T_3^3 = -e^{-\lambda}\left(\frac{\omega''}{2} + \frac{\omega'^2}{4} - \frac{\lambda'\omega'}{4}\right) + \frac{\ddot{\lambda}}{2} + \frac{\dot{\lambda}^2}{4} + \frac{\ddot{\omega}}{2} + \frac{\dot{\omega}^2}{4} + \frac{\dot{\lambda}\dot{\omega}}{4} - \Lambda = 0 \tag{1.17}$$

$$8\pi T_4^4 = e^{-\omega} - e^{-\lambda}\left(\omega'' + \frac{3}{4}\omega'^2 - \frac{\lambda'\omega'}{2}\right) + \frac{\dot{\omega}^2}{4} + \frac{\dot{\lambda}\dot{\omega}}{2} - \Lambda = 8\pi\rho \tag{1.18}$$

$$8\pi e^{\lambda} T_4^1 = -8\pi T_1^4 = \frac{\omega'\dot{\omega}}{2} - \frac{\dot{\lambda}\omega'}{2} + \dot{\omega}' = 0 \tag{1.19}$$

and all other T_ν^μ are identically equal to zero. Tolman's

notation has been used in equations (1.16) to (1.19) in which the dots represent partial differentiation with respect to t while the primes represent partial differentiation with respect to r.

The partial differential equations (1.16) to (1.19), if they can be solved, will determine the density ρ and the values of λ and ω at all times and places. Being partial differential equations, the solutions will involve a number of arbitrary functions of r or t. These arbitrary functions actually serve a useful purpose since they will allow the imposition of a certain number of boundary conditions on a given model, such as initial density distribution, initial particle velocities and accelerations. After this imposition of boundary conditions the subsequent temporal and spacial history of the model can be calculated. Unfortunately, the partial differential equations are non-linear. They cannot, therefore, be solved by straightforward methods, but only by artifice. We shall now examine a group of artifices by which a formal solution can be obtained in the most general case and more useable solutions can be obtained for certain physically interesting special cases.

Let us begin with the simplest equation (1.19). After rearranging this equation, we can write it

$$\dot{\lambda} - \dot{\omega} = 2\frac{\dot{\omega}'}{\omega'} \tag{1.20}$$

Each side of this equation can readily be integrated with respect to t. Carrying this out, we have

$$\lambda - \omega = 2 \log \omega' + \varphi(r) \tag{1.21}$$

where $\varphi(r)$ is an arbitrary and unknown function of r. The arbitrary function must be added to the integrand because partial, rather than complete, differentials are involved. The solution (1.21) is now a perfectly general relation between λ and ω and not simply a special case. Let us define a new arbitrary function of r as $4(g + 1) = e^{-\varphi(r)}$. Then we can rewrite (1.21) as exponentials in the form

$$e^{\lambda} = e^{\omega} \frac{\omega'^2}{4(g+1)} \tag{1.22}$$

This is not the same form as derived by Tolman (Ref. 2, eq. 8), but this definition of g as the arbitrary function of r will be found to be more convenient in later equations.

An examination of (1.22) will show that g can range in value only between -1 and $+\infty$. From their use in the line-element (1.14) e^{λ} and e^{ω} can be only positive and real. In order that e^{ω} have only positive real values, ω must be a real function of r and t. Consequently the square of the partial derivative ω'^2 can be only a positive real number. Therefore the denominator on the right-hand side of (1.22) can vary only through

real values. This places the stated limits on the otherwise perfectly arbitrary function g.

Having found an expression for e^{λ} in (1.22) we may substitute this into the line-element (1.14) which gives

$$ds^2 = -e^{\omega}\left[\frac{\omega'^2}{4(g+1)}\,dr^2 + d\theta^2 + \sin^2\theta\, d\phi^2\right] + dt^2 \tag{1.23}$$

This line-element (1.23) is the non-homogeneous analog of the well-known Friedman-Robertson non-static homogeneous line-element.

Further useful integrals can be obtained from the three remaining differential equations (1.16) to (1.18). Substituting (1.22) into (1.16) and rearranging, we obtain

$$e^{\omega}\left(\frac{3}{4}\dot{\omega}^2 - \Lambda + \ddot{\omega}\right) = g \tag{1.24}$$

Multiplying (1.24) by $\dot{\omega}\, e^{\frac{1}{2}\omega}$ and again rearranging, we see that the equation may be written as

$$\frac{\partial}{\partial t}\left[e^{\frac{3}{2}\omega}\left(\frac{\dot{\omega}^2}{2} - \frac{2}{3}\Lambda\right)\right] = 2g\,\frac{\partial}{\partial t}\,e^{\frac{1}{2}\omega} \tag{1.25}$$

Again both sides of this equation can readily be integrated with respect to t giving

$$e^{\frac{3}{2}\omega}\left(\frac{\dot{\omega}^2}{2} - \frac{2}{3}\Lambda\right) = 2g e^{\frac{1}{2}\omega} + 2h \tag{1.26}$$

where h is another arbitrary function of r necessary in the integrand to give a general solution.

However, equation (1.26) is still a partial differential equation. A further integration is necessary to express the temporal and spacial behavior of e^{ω} in analytic form. By multiplying (1.26) by $\frac{1}{2}e^{-\frac{1}{2}\omega}$ and re-arranging, this equation may be written as

$$\frac{1}{4}\dot{\omega}^2 e^{\omega} = g + he^{-\frac{1}{2}\omega} + \frac{1}{3}\Lambda e^{\omega} \tag{1.27}$$

This equation might appear more familiar if the dot notation is replaced by the more conventional partial differentiation signs as

$$\left(\frac{\partial}{\partial t} e^{\frac{1}{2}\omega}\right)^2 = g + he^{-\frac{1}{2}\omega} + \frac{1}{3}\Lambda e^{\omega} \tag{1.28}$$

The integration of (1.28) is the crux of the work which is to follow. It will be noticed that the two arbitrary functions of r that have been introduced, namely g and h, are both somewhat different from the similar functions used by Tolman (Ref. 2). The functions defined here have been found, however, to reduce the bulk of the expressions somewhat and have proved, therefore, more convenient for use.

There now remain two unused partial differential equations, namely (1.17) and (1.18). Substitution of (1.22) into (1.17) merely yields (1.24) again, and so nothing new is gained. Substitution of (1.22) into (1.18), however, will yield an expression for the density ρ in the model.

Using (1.20) and (1.22) in (1.18), we obtain

$$8\pi\rho = \frac{3}{4}\dot{\omega}^2 + \frac{\dot{\omega}\dot{\omega}'}{\omega'} - e^{-\omega}\left(g + \frac{2g'}{\omega'}\right) - \Lambda \tag{1.29}$$

This equation may be put into a more simple form by combining it with (1.26). Differentiating (1.26) with respect to r, we obtain

$$2\frac{dh}{dr} = e^{\frac{3}{2}\omega}\omega'\left[\frac{3}{4}\dot{\omega}^2 + \frac{\dot{\omega}\dot{\omega}'}{\omega'} - e^{-\omega}\left(g + \frac{2g'}{\omega'}\right) - \Lambda\right] \tag{1.30}$$

Hence, it is obvious that the density equation may be written in the particularly simple form

$$8\pi\rho = \frac{2e^{-\frac{3}{2}\omega}}{\omega'}\frac{dh}{dr} \tag{1.31}$$

Equation (1.31) shows that the otherwise arbitrary function h can never be a constant if these models are to contain matter.

We have now developed enough relationships that we may return to the discussion of allowable motions in these models. From their use in the line-element both e^{λ} and e^{ω} must be real and positive at all times and for all values of r. However, there are no restrictions as yet on the values. Mathematically $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ could be either positive or negative or could oscillate in sign as the variables r and t change. Physically, a stricter condition must be imposed.

Let us consider an instantaneous measurement of total proper distances from the origin of the co-moving coordinates to a particular co-moving coordinate point r_o, Θ_o, ϕ_o at a particular time t_o. We will arbitrarily make these measurements from the origin to the reference point on the r_o - sphere, which length we will denote as l_1; then along a longitude line of the r_o - sphere, we shall denote this length as l_2; finally along a latitude line on the r_o - sphere to the coordinate point, this length will be denoted as l_3. These proper lengths are related to the $g_{\mu\nu}$ for the particular time t_o and to the co-moving coordinate system by the equations

$$l_1 = \int_0^{r_o} e^{\frac{1}{2}\lambda}\, dr \qquad l_2 = \int_0^{\Theta_o} e^{\frac{1}{2}\omega}\, d\Theta \qquad l_3 = \int_0^{\phi_o} e^{\frac{1}{2}\omega} \sin\Theta\, d\phi \tag{1.32}$$

We can regard the l_1, l_2, and l_3 measures as constituting another coordinate system. Then the prohibition against "mixing" motions of the particles can be restated as a requirement that the mapping of the co-moving coordinates upon the l-measure coordinates must be strictly one-to-one at all times and places. If two or more distinct sets of co-moving coordinates, say (r_1, Θ_1, ϕ_1) and (r_2, Θ_2, ϕ_2), have identical values of l_1, l_2, l_3 at a given time, the physical meaning is that they all occupy the same space. Since these are non-static models, this implies that the neighborhoods of co-moving coordinate-1 and of co-moving

coordinate-2 are passing through each other. This, as discussed earlier, is a direct violation of our basic postulate of zero pressure everywhere within the model.

An examination of (1.32) will show that one-to-one mapping is possible only if the algebraic signs of $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ are invariant. In other words, $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ may be either negative or positive, but they cannot oscillate between the two signs. While negative values of $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ are permissible, they add nothing new to the physics of the situation. Negative values must be assigned to $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ if the co-moving coordinates r, Θ, and ϕ are stated as negative numbers themselves. In the work which is to follow we shall assume that the co-moving coordinates are enumerated in the same sense and sign as the 1-measure coordinates. Consequently, we shall restrict $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ solely to positive values.

The restriction of these terms to positive values in turn places definite conditions on the otherwise completely arbitrary functions g and h. An examination of (1.26), bearing in mind that $e^{\frac{3}{2}\omega}$, $e^{\frac{1}{2}\omega}$ and $\dot{\omega}^2/2$ are both positive and real while Λ and g have real values, will show that the arbitrary function h must be a real function.

The restriction of $e^{\frac{1}{2}\lambda}$ to positive values places conditions upon $2(g + 1)^{\frac{1}{2}}$ and on the r-dependence

of the function, ω . An examination of (1.21) will show that the arbitrary function $\varphi(r)$ may at times be complex in value. While ω' must be real, it is free to assume negative values. When ω' is negative, $2 \log \omega' = 2 \log|\omega'| \pm 2\pi i$. Consequently, $\varphi(r)$ must have an equal and opposite imaginary part, that is $\varphi(r) = \text{Real}\,\varphi(r) \mp 2\pi i$ because the left-hand part of (1.21) must, by definition, be real. This does not alter our previous conclusion that $4(g + 1) = e^{-\text{Real}\,\varphi(r) \pm 2\pi i}$ must be positive at all times, but it does allow $2(g + 1)^{\frac{1}{2}} = e^{-\frac{1}{2}\text{Real}\,\varphi(r) \pm \pi i}$ to assume either positive or negative values. Moreover, we see that ω' and $2(g + 1)^{\frac{1}{2}}$ must always have the same sign. These conclusions are substantiated if (1.22) is rewritten as

$$e^{\frac{1}{2}\lambda} = e^{\frac{1}{2}\omega} \frac{\omega'}{2\sqrt{g+1}} \tag{1.33}$$

when we recall that $e^{\frac{1}{2}\lambda}$ and $e^{\frac{1}{2}\omega}$ are both limited to positive values.

A physical condition that must be observed is that the density can never be negative. In the equation for density (1.31) the right-hand side is made up of three terms. The factor $e^{-\frac{3}{2}\omega}$ is always positive since it is the product of two positive terms. Consequently, ω' and $\frac{dh}{dr}$ must again both be of the same sign. The physical restrictions on the arbitrary functions can now be summerized as the requirement that $\frac{dh}{dr}$, $\frac{\partial \omega}{\partial r}$, and $2(g + 1)^{\frac{1}{2}}$ must always be of the same sign.

To conclude this chapter, let us return to (1.28), the integration of which under various conditions will make up most of the subject matter in the following chapters. In (1.26) the dependent variable is actually $e^{\frac{1}{2}\omega}$, and so the mathematical notation can be simplified by taking $y = e^{\frac{1}{2}\omega}$ which reduces (1.28) to

$$y\left(\frac{\partial y}{\partial t}\right)^2 = g y + h + \frac{1}{3}\Lambda y^3 \tag{1.34}$$

The first integral (1.22) assumes a particularly simple form in this notation

$$e^{\lambda} = \frac{1}{g+1}\left(\frac{\partial y}{\partial r}\right)^2 \tag{1.35}$$

While the density integral (1.31) is not greatly clarified by this substitution

$$8\pi\rho = 3\,\frac{\frac{dh}{dr}}{\frac{\partial}{\partial r}(y^3)} \tag{1.36}$$

By taking the square-root of both sides, we can write equation (1.34) as a linear partial differential equation of the first order in the standard form

$$P(r,t,y)\frac{\partial y}{\partial t} + Q(r,t,y)\frac{\partial y}{\partial r} = R(r,t,y) \tag{1.37}$$

where $P = \sqrt{y}$ $\quad Q = 0$ $\quad$ and $R = \sqrt{gy + h + \frac{1}{3}\Lambda y^3}$

Such equations are easily solved by the method of Lagrange.

The corresponding system of ordinary differential equations for the characteristics is

$$\frac{dt}{P} = \frac{dr}{Q} = \frac{dy}{R} \tag{1.38}$$

If any two independent solutions to (1.38) can be found, such as $u = a$ and $v = b$, then the general solution to (1.37) is any arbitrary function of these two solutions, such as $\phi(u,v)$. For the second term of (1.38), one solution is obviously $r = r_o$, where r_o is any positive number. Using this solution with the first and last terms of (1.38), we find the second solution will be the solution of

$$\frac{dt}{dy} = \frac{P(r_o,t,y)}{R(r_o,t,y)} \tag{1.39}$$

This is an ordinary differential equation since r is to be considered as a constant during its integration. Let us indicate its solution as $\psi(t,r_o,y) = b$. Then the general solution of (1.37) is $\psi(t,r,y) = f(r)$, where r is again a completely independent variable and $f(r)$ is an arbitrary function since this is an arbitrary function of u and v.

Following the method of Lagrange, the general solution to the partial differential equation (1.34) may be written as

$$t + f = \int \frac{\sqrt{y}\ dy}{\sqrt{gy + h + \frac{1}{3}\Lambda y^3}} \tag{1.40}$$

where f is an arbitrary function of r. The integral is to be evaluated according to (1.39) by taking the independent variable r to be constant ($r = r_o$). After the integral is formally evaluated, then the constant value r_o is to be replaced once more by the independent variable r.

The integral in (1.40) can be evaluated under the most general conditions in terms of elliptic functions, if a simple transformation is made in the independent variable. The integral can be evaluated in terms of better known analytical functions only in those special cases in which the cubic under the radical either is reduced in degree or has repeated roots. Moreover, it will be noted that the solution (1.40) after integration is an expression for t in terms of r and y. This is a result of (1.39) which can be integrated only in the form $t + b = \psi(r_o, y)$. It will be found that an explicit statement of y in terms of r and t in known functions is possible only under highly restricted special conditions.

However, it is y that is the variable of physical interest. If we cannot express y explicitly in terms of known functions, perhaps we can express it explicitly in some other manner, such as in terms of functions involving a parameter. It will be observed that (1.34) is suggestive of Weierstrass's or Jacobi's elliptic differential equation. If (1.34) can be altered by transformations into one of these

known differential equations, then y can be explicitly given in terms of the transformed independent variables. The necessary transformations must reduce the cross-term to a simple square of a derivative and must reorder the polynomial on the right-hand side into one of the normal forms.

Let us introduce a new independent variable ξ which is to be a function of both r and t. By the introduction of this new variable we wish to transform $y(\frac{\partial y}{\partial t})^2 = \phi(r,t,y)$ into $(\frac{\partial y}{\partial \xi})^2 = \phi(r,\xi,y)$. With the introduction of the new variable the derivative becomes $\frac{\partial y}{\partial t} = \frac{\partial y}{\partial \xi}\frac{\partial \xi}{\partial t}$. Now ξ was introduced under the most general conditions. Consequently, one condition may be imposed upon it to make ξ a definite function of r and t. For this condition let us take

$$\frac{\partial \xi}{\partial t} = \frac{K}{y} \tag{1.41}$$

where K is a function of r which we can specify to suit our convenience. Then (1.34) is transformed into

$$K^2\left(\frac{\partial y}{\partial \xi}\right)^2 = gy^2 + hy + \tfrac{1}{3}\Lambda y^4 \tag{1.42}$$

The polynomial in (1.42) may now be transformed into either Weierstrass's or Jacobi's normal form by a simple transformation of the dependent variable. The partial differential

equation (1.42) is readily solved by Lagrange's method. For this equation Lagrange's method is effectively one of replacing the partial derivative by an ordinary derivative and replacing the independent variable r by a constant value $(r = r_o)$ for purposes of integration. Then, under these conditions, a formal solution is found. In the resulting formal solution the constant r_o is once more replaced by the independent variable r for a solution to the partial differential equation (1.42).

The imposed condition (1.41) can also be integrated by Lagrange's method. This yields the equation

$$t + f = \int \frac{y\, d\xi}{K} \tag{1.43}$$

where f is the previous arbitrary function of r. As before, the integral in (1.43) is to be evaluated under the formal condition of $r = r_o$, after which it is again replaced by the independent variable r.

The pair of equations (1.42) and (1.43) are fully equivalent to the previous solution (1.40). If the solutions to (1.42) and (1.43) are found, they will be of the form $y = \phi(r, \xi)$ and $t + f = \psi(r, \xi)$. Thus this pair of solutions are parametric equations for y and t in terms of a parameter ξ and an independent variable r. The evident advantage of this procedure is the explicit

statement of y in terms of known functions. Moreover, in the work which follows it will be seen that the equation pair (1.42) and (1.43) is more elegant to handle than the more conventional equation (1.40).

Chapter II

General Solutions

In this chapter we shall consider the solution of the partial differential equation (1.34) under the most general conditions. Such general solutions can be expressed only in terms of elliptic functions. In following chapters we shall examine various non-elliptic solutions and approximations that can be made under certain special assumptions.

The nature of the possible general solutions can be obtained from the differential equation (1.34) without actually solving it. Let us rewrite this equation in the form

$$\left(\frac{\partial y}{\partial t}\right)^2 = \frac{1}{y}\left(\frac{1}{3}\Lambda y^3 + gy + h\right) \tag{2.1}$$

We know that $y(r,t)$ is restricted to real positive values, that $g(r)$ and $h(r)$ are real functions where only $g(r)$ is somewhat restricted, and that Λ is theoretically an unrestricted real constant. Actually, from the closeness of the Newtonian approximation in celestial mechanics, we know that Λ must be an extremely small number, if not actually zero. However, in general terms, we can say that the values of the variable and the coefficients in (2.1) must lie somewhere within the following ranges

$$-\infty \leqslant \Lambda \leqslant +\infty$$

$$0 \leqslant y(r,t) \leqslant +\infty$$

$$-1 \leqslant g(r) \leqslant +\infty$$

$$-\infty \leqslant h(r) \leqslant +\infty$$

Moreover we can say that $0 \leqslant \left(\frac{\partial y}{\partial t}\right)^2$ since a negative value for $\left(\frac{\partial y}{\partial t}\right)^2$ would require a complex value for $y(r,t)$.

For those solutions which include $y = 0$, we see that as $y \longrightarrow 0$ that $\left(\frac{\partial y}{\partial t}\right)^2 \approx h(r)/y$ and hence that $y(r,t) \approx \left(\frac{9}{4} h\right)^{1/3} (t+f)^{2/3}$ in the neighborhood of $y = 0$. We see, moreover, that the only solutions which will include $y = 0$ are those for which $h(r) \geqslant 0$. Every solution with $h(r) < 0$ must have a finite minimum value for $y(r,t)$.

On the other hand, for those solutions where $y(r,t)$ approaches positive infinity, we see that $\left(\frac{\partial y}{\partial t}\right)^2 \approx \frac{1}{3} \Lambda y^2$ as $y \longrightarrow +\infty$ and hence that in this neighborhood that $y(r,t) \approx e^{\sqrt{\frac{\Lambda}{3}}\,|t+f|}$. Consequently, the only solutions in which $y(r,t)$ can approach infinity are those having $\Lambda \geqslant 0$. Every solution having $\Lambda < 0$ must have a finite maximum value for $y(r,t)$.

Let us now consider the "turning points" of $y(r_o, t)$ as a function of t at a particular value of $r = r_o$. With the understood assumption that $y(r_o,t)$ is a continuous function, there are only two types of possible "turning

points". When $\left(\frac{\partial y}{\partial t}\right)^2 = +\infty$, the turning point is a cusp. From the nature of the right-hand side of (2.1), this is possible in the finite range only at $y = 0$. When $\left(\frac{\partial y}{\partial t}\right)^2 = 0$, the turning point is smooth. This turning point might be either a maximum, a minimum, or a point of inflexion. Moreover these smooth turning points can occur only at the roots of the cubic which is contained within the parenthesis in (2.1). This cubic can have a maximum of three real roots, of which only a maximum of two could be positive since the y^2 term in the cubic is zero. However, this cubic will have three real roots only if the value of its discriminant is less than zero. Stated mathematically, this condition is that $\left(\frac{\Lambda}{9}\right)^3\left[9^3 + \frac{9}{4}\Lambda h^2\right] < 0$. If the value of the discriminant is greater than zero, then the cubic can have only one real root, which might be either positive or negative in value. An allowable solution to (2.1) can have a point of inflexion only at a positive repeated root of the cubic, and only then if the cubic is positive for values of $y(r_o,t)$ which are both greater and less than the repeated root. This specialization produces a non-elliptic solution which will be considered in a later chapter. However, there it will be shown that the allowable solutions do not pass through the repeated root, but approach it asymptotically. Thus, the solutions to (2.1) do not have points of inflexion, but only maxima and minima,

varying monotonically between them.

Since we need a system to classify our allowable solutions, we shall adopt the notation first evolved by Robertson (Ref. 4, p. 74 et seq.) and extended by Tolman (Ref. 1, p.398 et seq.) for use with homogeneous models. We shall use the Robertson-Tolman notation to describe the local behavior of our non-static, non-homogeneous models. Thus, if an observer at r_o would determine that the behavior of matter in his immediate neighborhood was identical with what he would have expected in a homogeneous O_1 model, we shall classify the behavior of the model at r_o as being O_1. As this observer explored his surrounding space more deeply, he would discover, however, that the behavior of matter is different from his own at a distance, since the determining functions $g(r)$ and $h(r)$ are free to vary with radius vector. Let us determine, therefore, the types of behavior that might be expected at r_o with different values of $g(r_o)$ and $h(r_o)$.

First, let us find the types of solutions that could be expected when $\Lambda > 0$. We shall assume that $g(r_o)$, $h(r_o)$, and Λ are all finite and non-zero. When any of these three terms is allowed to become either zero or infinite, we shall find specialized non-elliptic solutions, which will be treated elsewhere. Let us break the problem up into quadrants according to the sign of $g(r_o)$ and $h(r_o)$.

1. $g(r_o) > 0 \qquad h(r_o) > 0$. Here the cubic, and hence $\left(\frac{\partial y}{\partial t}\right)^2$ is positive over the entire allowed range of $y(r_o,t)$. Therefore, $y(r_o, -\infty)$ is infinite, while $y(r_o,t)$ decreases monotonically until $y = 0$ at $t + f = 0$, where it goes through a cusp turning point after which it increases monotonically until $y(r_o, +\infty)$ is again infinite. This local behavior is identical with the model behavior of a monotonic universe of the first type in the Robertson-Tolman notation (M_1).

2. $g(r_o) < 0 \qquad h(r_o) > 0$. Here the cubic will have two real positive roots if the discriminant is less than zero. The condition that must be satisfied is that $|g^3| > \frac{9}{4} \Lambda h^2$. Let us indicate these two positive roots by E_1 and E_2, where $E_2 > E_1$. Then there are two independent ranges for $y(r_o,t)$ which will give $\left(\frac{\partial y}{\partial t}\right)^2 > 0$. Since the cubic is positive for $0 \leqslant y(r_o,t) \leqslant E_1$, this is an allowed solution with cusp turning points at $y = 0$ and smooth maximum turning points at $y = E_1$ while oscillating monotonically between these two values. This local behavior is equivalent to an oscillating universe of the first type and is denoted as O_1. The cubic is also positive for $E_2 \leqslant y(r_o,t)$. This solution has $y(r_o, -\infty) = +\infty$ decreasing monotonically until $y = E_2$ where it has a smooth minimum and then increasing monotonically until $y(r_o,+\infty) = +\infty$. This local behavior is equivalent to a monotonic universe

of the second type and will be denoted as M_2.

On the other hand, if the discriminant is greater than zero, the cubic will have no positive roots. The condition to be satisfied is, of course, that $|g^3| < \frac{9}{4} \Lambda h^2$. Here the cubic will be positive for all positive values of y and hence the allowable solution will again be similar to a monotonic universe of the first type and will be denoted as M_1.

3. $g(r_o) > 0 \qquad h(r_o) < 0$. In this case the cubic has one real positive root, which we may indicate as E. The cubic will be positive when $E \leqslant y(r_o,t) \leqslant +\infty$ and the allowed solution will again be locally equivalent to a monotonic universe of the second type and will be denoted as M_2.

4. $g(r_o) < 0 \qquad h(r_o) < 0$. The conditions here are similar to those in the preceding paragraph. The cubic has a single real positive root, say E. The cubic is positive when $E \leqslant y(r_o,t) \leqslant +\infty$ and the local behavior is similar to that of a monotonic universe of the second type and will be denoted as M_2.

Now let us repeat this procedure under the contrary assumption that $\Lambda < 0$ but that $g(r_o)$, $h(r_o)$, Λ are all finite and non-zero.

1. $g(r_o) > 0 \qquad h(r_o) > 0$. Here the cubic has a single real positive root, say E. The cubic is positive

when $0 \leq y(r_o,t) \leq E$. Hence, the local behavior is similar to an oscillating universe of the first type with cusp turning points at $y = 0$ and smooth maxima at $y = E$. This is the behavior which we have denoted as O_1.

2. $\underline{g(r_o) < 0 \qquad h(r_o) > 0}$. Again the cubic has a single real positive root at E. The cubic is positive in the domain $0 \leq y(r_o,t) \leq E$. The local behavior is similar to that of an oscillating universe of the first type and will be denoted as O_1.

3. $\underline{g(r_o) > 0 \qquad h(r_o) < 0}$. Once more the cubic will have two real positive roots, which we shall call E_1 and E_2 with $E_2 > E_1$ if the discriminant is less than zero. With the present signs for the three coefficients, the condition for two positive roots is that $g^3 > \frac{9}{4}|\Lambda| h^2$ Here, however, the cubic will be positive when $E_1 \leq y(r_o,t) \leq E_2$. The allowed solution will then have maxima at $y = E_2$ and minima at $y = E_1$ and will oscillate smoothly between these two values. This local behavior is similar to that of an oscillating universe of the second type and will accordingly be denoted as O_2.

Under the opposite condition of a discriminant greater in value than zero, the cubic will have no real positive roots. This condition is, of course, that $g^3 < \frac{9}{4}|\Lambda| h^2$. Here the cubic, and hence $\left(\frac{\partial y}{\partial t}\right)^2$, will be negative for all positive values of $y(r_o,t)$. Therefore

no physically possible solution exists under these conditions.

4. $g(r_o) < 0 \qquad h(r_o) < 0$. Since all of the coefficients of the cubic would be negative under these conditions, it is obvious that the cubic would itself be negative for all positive values of y. Hence it is again obvious that no physically possible solution could exist.

These allowable solutions and the conditions under which they exist are summarized in Table I. In this table a code letter has been assigned to each solution so that it can be identified in later discussion. Non-homogeneous models can be constructed having several different types of behavior since $g(r)$ and $h(r)$ are nearly arbitrary functions which are subject only to the condition that the resulting densities must not be negative at any place or time in the model. The density problem will be discussed in a later chapter after the actual elliptic solutions have been obtained. Complete freedom in mixing all types of behavior is impossible, of course, since the cosmological constant Λ is, by definition, constant throughout the model. Thus, if $\Lambda > 0$, then O_2 behavior is excluded. While if $\Lambda < 0$, then M_1 and M_2 behavior is forbidden.

Having determined the nature of all allowable solutions, let us now turn to the exact solution of the

Table I

Λ	g(r)	h(r)	Type	Range	Conditions	Code Letter
+	+	+	M_1	$0 \leq y \leq +\infty$		A
	−	+	O_1	$0 \leq y \leq E_1$	$\lvert g^3 \rvert > \frac{9}{4}\Lambda h^2$	B
			M_2	$E_2 \leq y \leq +\infty$		C
			M_1	$0 \leq y \leq +\infty$	$\lvert g^3 \rvert < \frac{9}{4}\Lambda h^2$	D
	+	−	M_2	$E \leq y \leq +\infty$		E
	−	−	M_2	$E \leq y \leq +\infty$		F
−	+	+	O_1	$0 \leq y \leq E$		G
	−	+	O_1	$0 \leq y \leq E$		H
	+	−	O_2	$E_1 \leq y \leq E_2$	$g^3 > \frac{9}{4}\lvert\Lambda\rvert h^2$	I
			No Solution		$g^3 < \frac{9}{4}\lvert\Lambda\rvert h^2$	
	−	−	No Solution			

of the partial differential equation (2.1). We shall consider first the solution pair (1.42) and (1.43). Let us transform (1.42) into Weierstrass's normal form and solve the resulting equation by Lagrange's method. The choice of Weierstrassian elliptic functions is quite arbitrary. The partial differential equation (1.42) can also be transformed into Jacobi's normal form and the solutions obtained in terms of Jacobian elliptic functions. Actually, extensive calculations have been made in both systems of elliptic functions. The Jacobian functions for the purposes of this work were found to be even more ungainly and difficult than the Weierstrassian elliptic functions and accordingly were discarded. None of the alternate theory in terms of Jacobian elliptic functions

will be given here.

Since the function $K(r)$ may be specified to suit our convenience, let us take $K(r) = (h/4)^{1/3}$. To transform the polynomial in (1.42) into Weierstrass's normal form, let us introduce a new dependent variable defined as $z = \frac{K(r)}{y(r,t)} + \alpha(r)$ where $\alpha(r) = \frac{g}{3(2h)^{2/3}}$ is solely a function of r. Then, for purposes of integration, (1.42) becomes effectively,

$$\left(\frac{dz}{d\xi}\right)^2 = 4z^3 - g_2 z - g_3 \tag{2.2}$$

where the invariants $g_2 = 12\alpha^2$ and $g_3 = -8\alpha^3 - \frac{1}{3}\Lambda$ are to be regarded as constants ($r = r_0$). If the three roots of the cubic in (2.2) are all distinct, the solution is well known to be the Weierstrassian elliptic function $z = \wp(\xi\,;\, g_2, g_3)$. Therefore, the first of the parametric equations will be

$$y(r,t) = \frac{K(r)}{\wp(\xi) - \alpha(r)} \tag{2.3}$$

where now the invariants $g_2(r)$ and $g_3(r)$ along with the functions $K(r)$ and $\alpha(r)$ are all taken to be functions of the independent variable r.

The cubic in (2.2) will have repeated roots for those special cases which will reduce its discriminant to zero. In these special cases the solution to (2.1) may be expressed

in terms of better known, non-elliptic functions, and so will be considered in the following chapters. The condition for a zero discriminant of this cubic is

$$\Lambda\left[\alpha^3 + \frac{1}{48}\Lambda\right] = 0 \tag{2.4}$$

Consequently, the two special cases $\Lambda = 0$ and $\alpha^3 = -\frac{1}{48}\Lambda$ are excluded from the discussion in this chapter.

The second parametric equation is found by substituting (2.3) into (1.43) giving

$$t + f(r) = \int \frac{d\xi}{\wp(\xi) - \alpha} \tag{2.5}$$

The elliptic function $\wp(\xi)$ is defined over the complex plane. Hence the parameter ξ may assume complex values. However, the parameter ξ is physically restricted to those values which will yield positive real values for $y(r,t)$ and which will give only real values for t. The consequences of this restriction will be discussed shortly.

Let us now return to our original solution (1.40) and show that its reduction leads also to equations (2.3) and (2.5). Thus we shall show the identity of a direct integration of (1.40) and of the more elegant use of the parametric equations (1.42) and (1.43). To integrate (1.40) let us invert and translate the dependent variable $y(r,t)$ into a new dependent variable $z(r,t)$, where z is defined

as $z = \frac{K(r)}{y} + \alpha(r)$ and $K(r)$ and $\alpha(r)$ are the same quantities which were used before. This reduces (1.40) into the transformed integral

$$t + f(r) = \int \frac{dz}{(z-\alpha)\sqrt{4z^3 - g_2 z - g_3}} \tag{2.6}$$

where g_2 and g_3 have been defined earlier. The integral (2.6) is an elliptic integral of the third kind. It is, in fact, Weierstrass's third normal form. This integral may be integrated by introducing Weierstrass's elliptic function. Taking advantage of the well known relationship for Weierstrass's elliptic function $\frac{d}{d\xi}\wp(\xi) = \sqrt{4\wp^3(\xi) - g_2\wp(\xi) - g_3}$ we make the additional transformation of $z = \wp(\xi\ ; g_2, g_3)$ which transforms the integral (2.6) into

$$t + f(r) = \int \frac{d\xi}{\wp(\xi) - \alpha} \tag{2.5}$$

Since this last integral is identical with one obtained from the parametric equations, it has been numbered to correspond.

Two transformations, one real and one elliptic, have been made in the dependent variable. Solving these transformations for y we have

$$y = \frac{K(r)}{\wp(\xi) - \alpha(r)} \tag{2.3}$$

Since this expression is again identical with that obtained from the first parametric equation, it has also been numbered to correspond. Thus a direct integration of (1.40) yields exactly the same two equations (2.3) and (2.5) as did the pair of parametric equations. Hence the two procedures for solution are perfectly equivalent.

The parametric equations (1.42) and (1.43) were deduced with great generality, but perhaps with a corresponding loss of intuitive reality. However, it is easily shown that the parametric equations are inherent in the straight-forward solution of (1.40), which has just been presented. In reducing (1.40) to (2.5) two transformations were made in the dependent variable. The transformation $z = \wp(\xi; g_2, g_3)$ requires that z satisfy the differential equation $\left(\frac{dz}{d\xi}\right)^2 = 4z^3 - g_2 z - g_3$ when r is held constant ($r = r_0$). When the other transformation $z = \frac{K(r)}{y} + \alpha(r)$ is substituted into this differential equation and solved for y, we have

$$K^2(r)\left\{\frac{\partial y(r,t)}{\partial \xi}\right\}^2 = g(r)\,y^2 + h(r)\,y + \tfrac{1}{3}\Lambda\,y^4 \qquad (1.42)$$

when we allow r to be an independent variable again. This is the first of the two parametric equations. The second parametric equation follows readily from the substitution of (2.3) into (2.5) which gives

$$t + f(r) = \int \frac{y\, d\xi}{K(r)} \tag{1.43}$$

Thus the parametric equations may be directly derived from the general solution (1.40).

The evaluation of the integral in (2.5) now remains to be done. The integrand in (2.5) is an elliptic function of the second order with a double zero at $\xi = 0$ and all congruent points and with simple poles at $\xi = \chi$ and $\xi = -\chi$ and all congruent points, where the constant χ is any one of the complex numbers which will satisfy $\wp(\chi) = \alpha$. Consequently, the integrand of (2.5) may be expressed in terms of Weierstrass's Zeta functions (Ref. 5, p. 369). Since the residue at the pole $\xi = \chi$ is $+\frac{1}{\wp'(\chi)}$ while the residue at the other pole $\xi = -\chi$ is $-\frac{1}{\wp'(\chi)}$, the integrand may be expressed as

$$\frac{1}{\wp(\xi) - \alpha} = \frac{1}{\wp'(\chi)} \left[-\zeta(\chi - \xi) - \zeta(\chi + \xi) + 2\zeta(\chi) \right] \tag{2.7}$$

where the constant term in the expansion by Zeta functions has been determined by the presence of a double zero at the origin.

Since Zeta functions are readily integrable in terms of Weierstrassian Sigma functions, the integral (2.5) becomes

$$t + f(r) = \frac{1}{\wp'(\chi)} \left[\log \frac{\sigma(\chi - \xi)}{\sigma(\chi + \xi)} + 2\xi\zeta(\chi) - F(\xi_0) \right] \tag{2.8}$$

The second constant of integration $F(\xi_0)$ is actually redundant since it could have been combined with the other

constant of integration $f(r)$. The redundant constant of integration $F(\xi_o)$ has been introduced to allow the zero-point of $t + f(r)$ to be adjusted at will with respect to ξ. If we should take $t + f(r) = 0$ when $\xi = 0$, then $F(0) = 0$. But if we wish to take the zero-point of $t + f(r)$ at one of the half-periods of $\wp(\xi)$, usually denoted as ω_1, ω_2, and ω_3, then $F(\omega_j) = 2\omega_j\zeta(x) - 2x\zeta(\omega_j) \pm \pi i$ where j may be either 1, 2, or 3. We shall see that this might be desirable for some of the solutions outlined in Table I.

The pair of equations (2.3) and (2.8) constitute a complete parametric solution to the differential equation (2.1). The solutions of (2.1) involving only non-elliptic functions which arise when the cubic in (2.2) has repeated roots are only special cases of the general solution given in (2.3) and (2.8). These special solutions may be found from the known degenerations of the elliptic functions in (2.3) and (2.8) when two or more of the roots are identical. The statement of the value of the elliptic integral given in (2.8) is not the one given in standard compendia such as Jahnke and Emde (Ref. 6, p.105). The form given in (2.8) was adopted to avoid difficulties in passing to the limit as two or more of the roots in the cubic in (2.2) approach a common value.

The elliptic functions in (2.3) and (2.8) are not convenient functions for actual calculation. No complete

tables exist for these functions. While there are numerical procedures by which the values of these elliptic functions may be obtained, they are so difficult that any extended calculation of $y(r,t)$ would hardly be feasible. In the following chapter some non-elliptic approximations to (2.3) and (2.8) will be given for the physically interesting case of $|\Lambda| < |\alpha^3|$. However, there are no easy calculational methods when Λ is comparable to α^3 in value. The general solution given by equations (2.3) and (2.8) must contain the four types of behavior diagrammed in Table I as well as all forms of special behavior to be discussed in later chapters. This separation of physical reality from the esoteric mathematical formalism of the general solution must be done from the known properties of the elliptic functions and from the permissible ranges of the parameter ξ as determined by physical considerations.

The elliptic function $\wp(\xi)$ is doubly-periodic with the two periods $2\omega_1$ and $2\omega_2$, whose ratio cannot be real. That is $\wp(\xi + 2m\omega_1 + 2n\omega_2) = \wp(\xi)$ when m and n are any positive or negative integers. This might be compared with a familiar singly-periodic function such as the trigonometric sine function which has a single period of 2π. This means that $\mathrm{Sin}\,(\xi + 2n\pi) = \mathrm{Sin}\,\xi$,

where again n is any integer. The single periodicity of $\mathrm{Sin}\,\xi$ means that it can be defined over the entire complex ξ-plane by defining $\mathrm{Sin}\,\xi$ within some fundamental period-strip such as $0 \leq \mathrm{Real}\,\xi \leq 2\pi$. The extension to the entire ξ-plane is then made by repeated duplications of the fundamental period-strip under successive translations of one period. In the same way the elliptic function $\wp(\xi)$ may be defined over the entire ξ-plane by defining it within a fundamental period-parallelogram. One possible choice for the fundamental period-parallelogram would be that parallelogram in the ξ-plane whose apices are $2\omega_1$, $2\omega_1 + 2\omega_2$, $2\omega_2$, and 0. Then $\wp(\xi)$ is defined over the entire ξ-plane by repeated duplications of this fundamental period-parallelogram under successive translations of the form $2m\omega_1 + 2n\omega_2$ where m and n compass all of the integers. Thus our study of equation (2.3) may be confined to a single period-parallelogram in the ξ-plane since it will merely repeat itself according to a regular pattern over the remainder of the ξ-plane.

It is well known that $\mathrm{Sin}\,\xi$ has poles at the extremities of its period-strip as $\mathrm{Imaginary}\,\xi \to \pm\infty$. Thus it is not surprising that the Weierstrassian elliptic function $\wp(\xi)$ has double poles at the apices of the fundamental period-parallelogram just outlined, namely at

0, $2\omega_1$, $2\omega_1 + 2\omega_2$, and $2\omega_2$. These poles repeat themselves, of course, over the entire ξ-plane according to the regular pattern. The relationships of the zeros within the period-strip of a singly-periodic function and within the period-parallelogram of a doubly-periodic function are also analogous. Thus a doubly-periodic function such as the Weierstrassian elliptic function might be thought of as a logical higher generalization of the intuitively more acceptable singly-periodic functions.

The Weierstrassian Sigma and Zeta functions used in equation (2.8) may be defined as $\zeta(u) = -\int \wp(u)\, du$ and $\log \sigma(u) = \int \zeta(u)\, du$. It should be noted that the Sigma and Zeta functions are not doubly-periodic but vary in a regular way from one period-parallelogram to another. Thus a study of equation (2.8) cannot be confined to a study of its behavior within a single period-parallelogram. However, in the only case in which the parameter ξ extends beyond a single period-parallelogram, namely, in the O_1 and O_2 solutions, it will be seen that the parameter ξ simply initiates a new cycle of oscillating behavior when it enters the adjacent period-parallelogram being equivalent, therefore, to a simple translation of the time axis. Further information on the Weierstrassian elliptic functions may be found in any of the standard works on mathematical analysis

(e.g., Ref. 5, Chap. 13). Notation for these functions has not been standarized but varies from author to author.

The periods of the elliptic function $\wp(\xi)$ are related to the three roots of the cubic in (2.2), which we shall denote by e_1, e_2, and e_3 by the relations

$$e_1 = \wp(\omega_1) \qquad e_2 = \wp(\omega_2) \qquad e_3 = \wp(\omega_3) \tag{2.9}$$

$$e_1 + e_2 + e_3 = 0 \qquad \omega_1 + \omega_2 + \omega_3 = 0$$

For a given set of three roots, six different elliptic functions may be formed from the six possible ways in which the two periods $2\omega_1$ and $2\omega_2$ may be associated with the three roots. Actually, as far as calculations are concerned, these six elliptic functions are perfectly equivalent. For a given value of ξ each of these six elliptic functions will yield exactly the same numerical value. The six seemingly different elliptic functions arise from the six possible choices for the fundamental period-parallelogram in the complex ξ -plane. For convenience in the work to follow, we need to settle upon one of the six possible representations as the one which will be used.

Each elliptic function $\wp(\xi; g_2, g_3)$ having real values for g_2 and g_3 has a possible fundamental period which is real. Let us take this real period to be positive and denote it as $2\omega_1$. Let us take the root e_1 as

associated with this real period by the relation $\wp(\omega_1) = e_1$. We shall see later that the root e_1 is real. We shall chose $2\omega_2$ so that the period-parallelogram lies within the first quadrant and so that $\mathrm{Arg}\,\omega_2 > \mathrm{Arg}(-\omega_3)$. With the period $2\omega_2$ will be associated the root e_2 by the relation $\wp(\omega_2) = e_2$. If all three roots are real, then $2\omega_2$ is a pure imaginary and $e_1 > e_3 > e_2$. If the two roots e_2 and e_3 are complex, then the period $2\omega_2$ is also complex. The period $2\omega_3 = -2\omega_1 - 2\omega_2$ is always complex.

The elliptic function $\wp(\xi)$ is physically restricted to real values, since $y(r,t)$, $K(r)$, and $\alpha(r)$ are all real quantities. Consequently, ξ is not free to vary over the entire ξ-plane but is confined to a family of allowable curves over the ξ-plane. These allowable curves may be delineated from a study of the inverse elliptic integral (vide Ref. 7, p.77) corresponding to the elliptic function $z = \wp(\xi; g_2, g_3)$, which is

$$\xi = \int_z^\infty \frac{du}{\sqrt{S}} \qquad \text{where} \quad S = 4u^3 - g_2 u - g_3 \,. \tag{2.10}$$

Because of the presence of the three poles of the integrand at the points e_1, e_2, and e_3 in the u-plane, the value of the integral (2.10) ξ is not independent of the path of integration. There is, in fact, a doubly infinite set of values ξ which correspond to a given value of z

depending upon how the path of integration loops around the poles at e_1, e_2, and e_3. This is, of course, nothing but the familiar statement of double-periodicity of the elliptic function $z = \wp(\xi)$. The mapping of the z-plane upon the ξ-plane can be made one-to-one only by introducing cuts and Riemannian surfaces into the u-plane.

Since the allowable values of ξ will depend upon whether the three roots e_1, e_2, and e_3 are all real or whether two of these roots are complex, the two cases will be considered separately. Let us first examine the conditions when e_1 is real and e_2 and e_3 are complex conjugates. From the discriminant of the cubic, this will be true whenever

$$\Lambda\left[\alpha^3 + \frac{1}{48}\Lambda\right] > 0 \qquad (2.11)$$

Under these conditions S is real and positive whenever $u > e_1$ and real, while S is real and negative whenever $u < e_1$ and real.

Let us now determine the range of ξ corresponding to z by integral (2.10) when $z \geqslant e_1$. Let us assume that the u-plane has been cut from each of the branch points e_1, e_2, and e_3 out to infinity. Let us take our contour for integration along the positive real axis in the u-plane. Along this contour, for $z > e_1$, S is everywhere positive.

Let us assume in the particular Riemann sheet in which our contour is situated that $S^{-\frac{1}{2}}$ is positive. Then it is obvious that $0 \leqslant \xi \leqslant \omega_1$ when $+\infty \geqslant z \geqslant e_1$. If other contours are chosen which loop around the three branch points in various ways, the range of ξ corresponding to $z \geqslant e_1$ is extended. Alternately, the complete possible range for ξ when $z \geqslant e_1$ can be found from the properties of $\wp(\xi)$, such as its evenness and its double-periodicity. The lines in the ξ-plane which delineate the possible values of ξ when $z \geqslant e_1$ and real will be found to consist of the real axis and all lines parallel to the real axis and passing through the poles of $z = \wp(\xi)$.

The values that ξ will assume when $z \leqslant e_1$ may be obtained in the same fashion. Since e_1 is the only real root, the other two being complex, S will be negative for all real u which are smaller than e_1. Consequently, $S^{-\frac{1}{2}}$ will be a pure imaginary along the real axis for all real $u < e_1$. Let us assume in the particular Riemannian sheet in which we have taken our contour that $S^{-\frac{1}{2}}$ is a positive imaginary. There will be at least one sheet in which this condition and the one of the preceding paragraph are valid. Hence, for the simple contour along the real axis and making an infinitely small half-circle around the branch point at e_1, we see by (2.10) that $\xi = \omega_1 + i\,\varphi(z)$ (function of z) when $z \leqslant e_1$. It is easy to show that

$\xi \rightarrow 2\omega_2$ when $z \rightarrow -\infty$. We can prove this in the limit by taking the closed contour along a large segment of the real axis, a large semi-circle, and completed by a small contour around the branch point e_2 and long contours along each side of the cut to e_2 . Thus, for this particular Riemannian sheet, we have $\omega_1 \leqslant \xi \leqslant 2\omega_2$ when $z \leqslant e_1$. This puts a definite restriction upon the proportions of the period-parallelogram since it requires that Real $2\omega_2 = \omega_1$. This requirement will be supported by the approximate solutions in the following chapter. By using other contours or by using the known properties of $\wp(\xi)$, we see that the lines in the ξ -plane which delineate the possible values of ξ when $z \leqslant e_1$ and real consist of the imaginary axis and all lines parallel to the imaginary axis and passing through the poles of $z = \wp(\xi)$.

The situation which has been discussed is diagrammed in Fig. 2.1. Here part of the ξ-plane is shown with the lines which contain all of the values of ξ which will yield real values for $z = \wp(\xi)$. The solid horizontal lines contain all of those ξ which will give real values of $z \geqslant e_1$, while the dotted vertical lines contain all of the ξ which will give real values of $z \leqslant e_1$. The diagonal broken lines have been ruled in purely as an aid to the eye in visualizing the net of period-parallelograms into which the ξ-plane is divided. Actually, only the horizontal

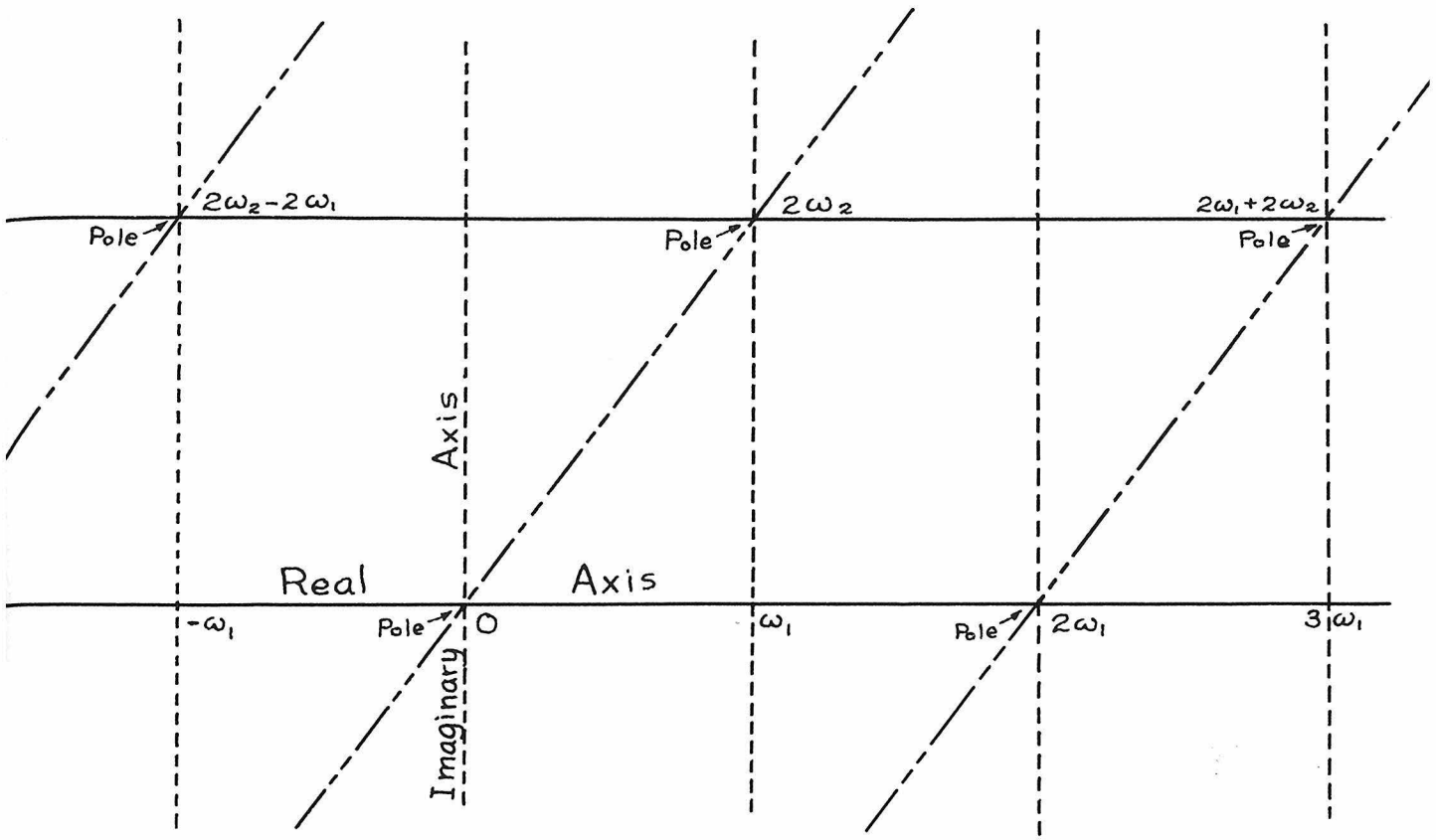

Fig. 2.1

segment from 0 to ω_1 and the vertical segment from ω_1 to $2\omega_2$ have been discussed. The remainder of the allowed line system has been derived either from a consideration of alternative contours of integration or from the properties of $\wp(\xi)$. We have also proved en passant by (2.9) and (2.10) that if e_1 is taken as the one real root, then the period $2\omega_1$ is real and positive under the conditions that we have assumed.

Geometrically the system of horizontal and vertical lines must intersect. There are only two classes

of points at which they can intersect without creating an ambiguity in value. One such class of points is $\xi = \omega_1$ and all congruent points since at these points $z = e_1$. The remaining class of acceptable points of intersection is at the poles of $\wp(\xi)$ since no ambiguity in value can be said to exist at a singularity. But from the theory of $\wp(\xi)$ we know that $\text{Arg}\left(\frac{\omega_2}{\omega_1}\right) < \frac{\pi}{2}$ when e_2 and e_3 are complex. Thus the requirement that Real $2\omega_2 = \omega_1$ is again verified.

Let us now consider the second case in which the three roots e_1, e_2, and e_3 of the cubic in (2.2) are all real. From the discriminant of the cubic, this will be true whenever

$$\Lambda\left[\alpha^3 + \frac{1}{48}\Lambda\right] < 0 \qquad (2.12)$$

Let us number our roots in the following order: $e_1 > e_3 > e_2$. Then the quantity S in (2.10) will have the following signs

$+\infty > u > e_1$	$S > 0$
$e_1 > u > e_3$	$S < 0$
$e_3 > u > e_2$	$S > 0$
$e_2 > u > -\infty$	$S < 0$

Let us once more assume that the u-plane is cut and multi-sheeted. Let us take a simple contour of integration for (2.10) along the real axis, slightly indented where necessary to pass any of the singularities at e_1, e_2, and e_3.

For the purpose of our discussion, let us choose the particular Riemann sheet in which

$+\infty > u > e_1$	$S^{-\frac{1}{2}}$	positive real
$e_1 > u > e_3$	$S^{-\frac{1}{2}}$	positive imaginary
$e_3 > u > e_2$	$S^{-\frac{1}{2}}$	negative real
$e_2 > u > -\infty$	$S^{-\frac{1}{2}}$	negative imaginary

Then by the previous reasoning, we would have

$+\infty \geqslant z \geqslant e_1$	$0 \leqslant \xi \leqslant \omega_1$
$e_1 \geqslant z \geqslant e_3$	$\omega_1 \leqslant \xi \leqslant \omega_1 + \omega_2$
$e_3 \geqslant z \geqslant e_2$	$\omega_1 + \omega_2 \geqslant \xi \geqslant \omega_2$
$e_2 \geqslant z \geqslant -\infty$	$\omega_2 \geqslant \xi \geqslant 0$

The last statement that $\xi \longrightarrow 0$ when $z \longrightarrow -\infty$ may be shown in the limit from a contour that is made up of a large segment of the real axis and completed by a large semi-circle on the side away from the cuts. Since the closed contour does not contain any singularities, the foregoing limit can be proved.

The foregoing ranges of ξ for real values of $z = \wp(\xi)$ may be extended, as previously, either by the consideration of alternate contours or by the properties of $\wp(\xi)$. The situation for three real roots is shown in Fig. 2.2. This plotted portion of the ξ-plane shows segments of four families of allowed curves for the variation of ξ. These encompass all values of ξ

which will yield real values for $z = \wp(\xi)$ plotted according to the following scheme

$+\infty \geqslant z \geqslant e_1$	solid horizontal lines
$e_1 \geqslant z \geqslant e_3$	dotted vertical lines
$e_3 \geqslant z \geqslant e_2$	dotted horizontal lines
$e_2 \geqslant z \geqslant -\infty$	solid vertical lines

The system of solid vertical and horizontal lines outline the meshes of period-parallelograms. Again it is obvious that under our assumptions the period $2\omega_1$ is real and positive. With a little study it is also obvious that in this case we have the period $2\omega_2$ as a positive pure imaginary.

Any of the allowed lines for ξ in Fig. 2.1 or Fig. 212 will give an acceptable value for $y(r,t)$ in (2.3) if the sign of $K(r)$, which is determined solely by the sign of $h(r)$, is properly adjusted. However, the allowable values of ξ are further restricted by the parametric solution for the time t. The time t is always a real number and must vary through purely real numbers as ξ varies. The integrand of (2.5), which is $\frac{y(r,t)}{K(r)}$, will be a real number whenever ξ ranges over any of the families of lines shown in Fig. 2.1 or Fig. 2.2. Hence, if the path of integration for (2.5) contains a segment of any of the allowable lines which are parallel to the imaginary axis, then in that segment we would have $d\xi =$ (a pure imaginary).

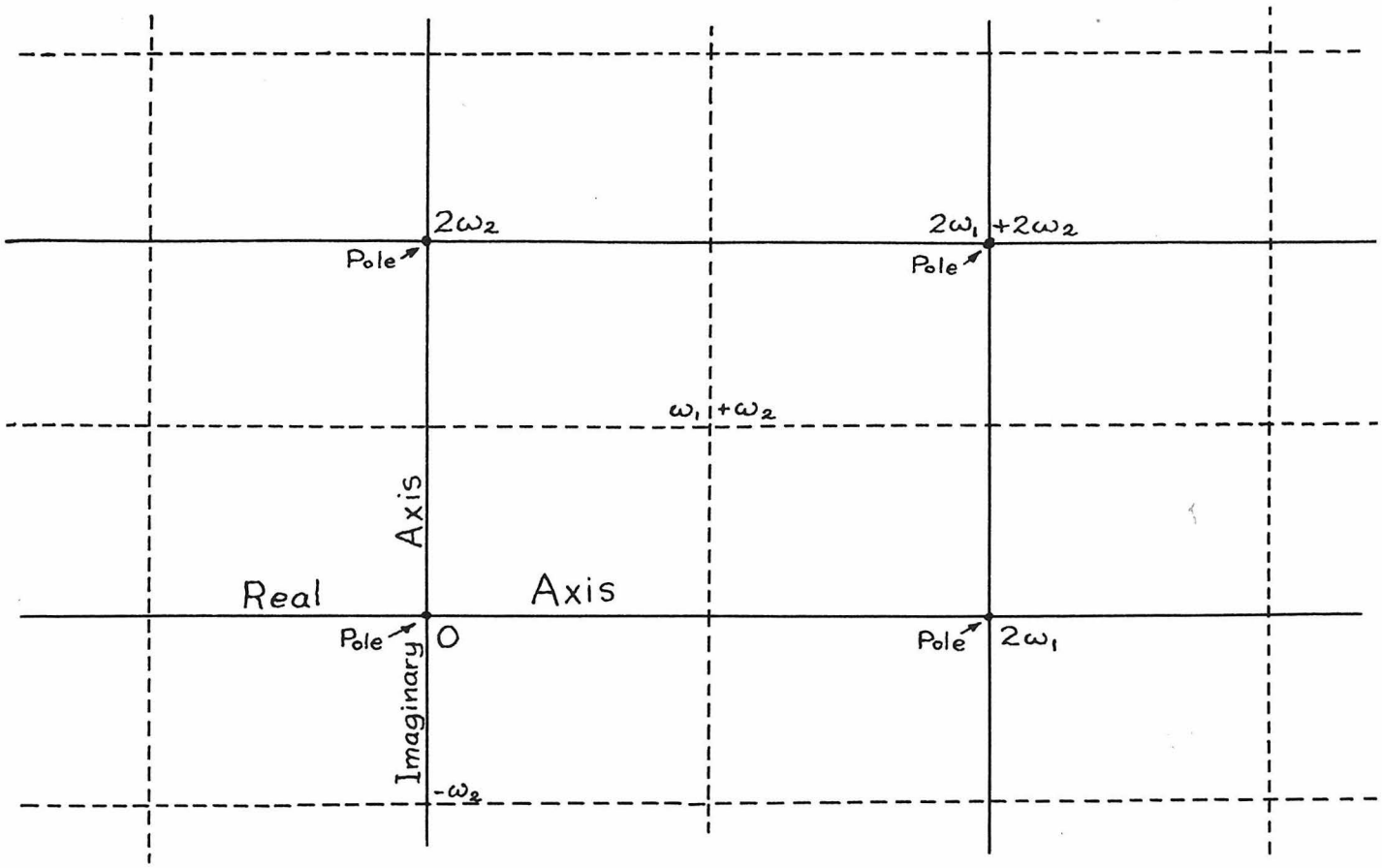

Fig. 2.2

Consequently, the value of the integral, which is $t + f(r)$, would have a purely imaginary increment when ξ ranges over that vertical segment. Since $f(r)$ is a constant for a given value of r, this imaginary increment is impossible. Therefore the path of integration of (2.5) cannot include any part of the imaginary axis or of any line parallel to it. Thus we see that the only physically allowable values of are those which are included within the horizontal families

of lines shown in Fig. 2.1 and Fig. 2.2. The vertical families of lines shown in these two figures, while acceptable for $y(r,t)$, must be excluded as yielding physically impossible values for t.

We have found valid ranges for the parameter ξ which will give solutions of real physical meaning to the general solutions (2.3) and (2.8). One valid range for the parameter ξ is the system of lines in the ξ-plane consisting of the real axis and all congruent lines resulting from the periodicity of $\wp(\xi)$. These congruent lines are also parallel to the real axis and displaced from it by $2n\omega_2$ where n is any positive or negative integer. While ξ varies along any line of this family, $+\infty \geqslant z = \wp(\xi) \geqslant e_1$, with $z = +\infty$ at $\xi = 0$ and all congruent points and with $z = e_1$ at $\xi = \omega_1$ and all congruent points. If the three roots are all real, a second valid range of ξ is possible. This consists of a line through ω_2 parallel to the real axis and all congruent lines produced by a displacement of $2n\omega_2$ from it. When ξ varies along any line of this second family, $e_2 \leqslant z = \wp(\xi) \leqslant e_3$, with $z = e_2$ at $\xi = \omega_2$ and all congruent points and with $z = e_3$ at $\xi = \omega_3$ and all congruent points. For a given value of Λ the roots are solely functions of the value of $\alpha(r)$. For the most general case the value of $\alpha(r)$ will vary with r and

consequently the behavior of $y(r,t)$ will also vary.

Therefore the relation of $\alpha(r)$ to the roots is important. In particular, if χ should lie upon one of the allowed lines for ξ in the ξ-plane, $y(r,t)$ would have a simple pole at $\xi = \chi$ and all congruent points. This, of course, is required for M_1 and M_2 solutions. Let us substitute $\alpha(r)$ into the cubic

$$S(\alpha) = 4\alpha^3 - g_2\alpha - g_3 = \tfrac{1}{3}\Lambda \qquad (2.13)$$

Therefore, while $\alpha(r)$ is free to vary through all of the real numbers, it will never cross a real root of the cubic since $S(\alpha)$ is a constant, independent of $\alpha(r)$. Hence if $\alpha(r)$ is located by any means as being between two real roots of the cubic, it will remain between those two real roots throughout its variation. In the case that the cubic possesses only one real root e_1 then if $\alpha(r)$ is established by any means as being greater or less in value than e_1, it will remain so throughout its variation.

Such determinations can be made in the limiting cases in which the cubic has repeated roots. In the limiting case that $\alpha(r) \to \pm\infty$, $\tfrac{1}{3}\Lambda$ can be neglected in g_3 as compared with $8\alpha^3$ and the cubic approaches

$$S(\alpha \to \pm\infty) = 4z^3 - g_2 z - g_3 \approx 4(z+2\alpha)(z-\alpha)^2 \qquad (2.14)$$

where the cubic has a real root of -2α and a real repeated root of α. In the other limiting case in which $\alpha^3 = -\frac{1}{48}\Lambda$ we have

$$S(\alpha^3 = -\tfrac{1}{48}\Lambda) = 4z^3 - g_2 z - g_3 = 4(z - 2\alpha)(z + \alpha)^2 \qquad (2.15)$$

in which the cubic has a real root of 2α and a real repeated root of $-\alpha$. The location of $\alpha(r)$ with respect to the roots from (2.13, (2.14), and (2.15) is given in Table II. Since we have arbitrarily taken e_1 to be either the only real root or to be the largest positive root, it has been necessary to reorder the roots in one instance.

After these rather lengthy preliminaries, we are now ready to particularize the general solution of equations (2.3) and (2.8). In the early part of this chapter we deduced certain conclusions directly from equation (2.1) as to the local behavior of the solutions under certain specified conditions. These conclusions are tabulated in Table I. We also found that if the solutions contain the values $y(r,t) = 0$ or $y(r,t) = +\infty$ that they must behave in a certain fashion in the neighborhood of these values. We will now examine the general solution to see if these conditions are fulfilled.

An inspection of Table II shows that the problem falls into four groups. Let us take these groups in order.

Table II

Λ	α	e_1	e_3	e_2	Range
+	$+\infty$	-2α	α	α	$\alpha > e_1$
	$\alpha^3 > -\frac{1}{48}\Lambda$	Real	Complex	Complex	
	$\alpha^3 = -\frac{1}{48}\Lambda$	2α	$-\alpha$	$-\alpha$	
	Change in notation of roots				
	$\alpha^3 = -\frac{1}{48}\Lambda$	$-\alpha$	$-\alpha$	2α	$e_2 \leqslant \alpha \leqslant e_3$
	$\alpha^3 < -\frac{1}{48}\Lambda$	Real	Real	Real	
	$-\infty$	-2α	α	α	
−	$+\infty$	α	α	-2α	$e_3 \leqslant \alpha \leqslant e_1$
	$\alpha^3 > -\frac{1}{48}\Lambda$	Real	Real	Real	
	$\alpha^3 = -\frac{1}{48}\Lambda$	2α	$-\alpha$	$-\alpha$	
	$\alpha^3 < -\frac{1}{48}\Lambda$	Real	Complex	Complex	$\alpha < e_1$
	$-\infty$	-2α	α	α	

First, let us consider the case in which $\Lambda > 0$ and $\alpha^3 > -\frac{1}{48}\Lambda$. Here we will have one real and two complex roots for the cubic in equation (2.2), and consequently the system of allowed lines for the variation of ξ shown in Fig. 2.1 will apply. Since $\alpha(r_0)$ is real and greater than e_1 in value, χ and its congruent points will lie on the allowed lines for ξ. For the sake of definiteness, let us take $0 < \chi < \omega_1$ where χ

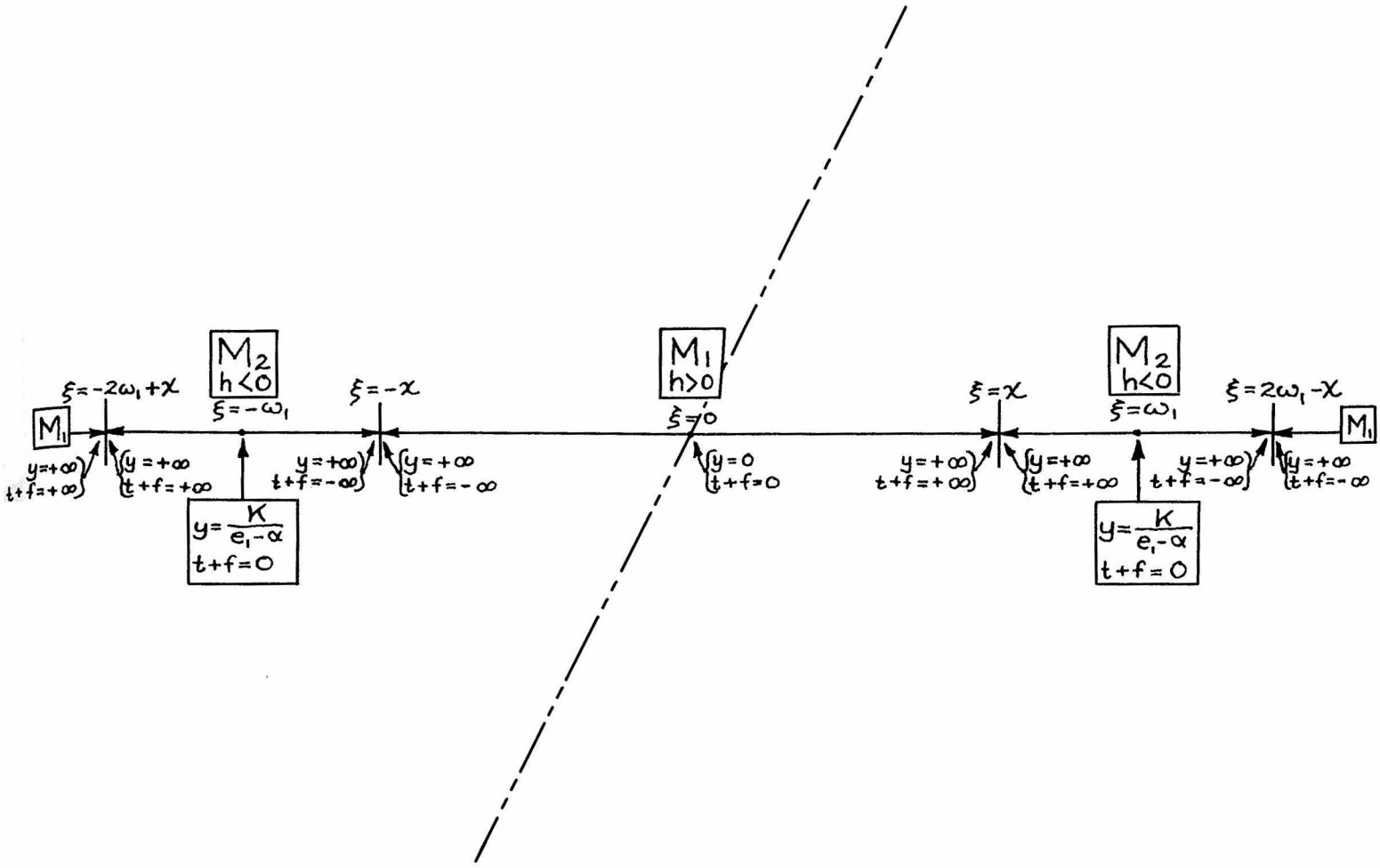

Fig. 2.3

is the real number which satisfies $\wp(\chi) = \alpha(r_0)$. Under these conditions, the situation in the ξ-plane is shown in Fig. 2.3. Only a part of one period-parallelogram is shown, but this is sufficient since these relations are repeated indefinitely over the mesh of the period-parallelograms.

There are two classes of solutions. When $-\chi \leqslant \xi \leqslant +\chi$ and in congruent segments, we have M_1 solutions. When $2\omega_1 - \chi \geqslant \xi \geqslant +\chi$ and in all congruent segments,

we have M_2 solutions. When $-\chi \leqslant \xi \leqslant +\chi$ and in all congruent segments, then $\wp(\xi) \geqslant \alpha(r_o)$; hence $K(r_o)$ in equation (2.3) must be positive. From the definition of $K(r)$, $h(r_o) > 0$ for the M_1 solutions. If $F(\xi_o) = 0$ in equation (2.8), then at $\xi = 0$ and congruent points, we have $y = 0$ and $t + f(r_o) = 0$. If equations (2.3) and (2.8) are expanded about the point $\xi = 0$ or about any of the congruent points and only the first term of the expansion retained, then we find $y \approx K(r_o)\xi^2$ and $t + f(r_o) \approx \frac{1}{3}\xi^3$. These two parametric equations may be combined into the explicit expression $y(r,t) \approx (9h/4)^{1/3}(t + f)^{2/3}$ which is valid in the neighborhood of $y = 0$. This is, of course, exactly the behavior predicted in the preliminary discussion for a solution in the neighborhood of $y = 0$. Moreover, since equation (2.3) can be zero only at $\xi = 0$ and congruent points (i.e., the poles of $\wp(\xi)$), this behavior in the neighborhood of $y = 0$ must be true for all possible valid O_1 and M_1 solutions.

Let us now examine the behavior of the M_1 solution in the neighborhood of $\xi = +\chi$. At the point $\xi = +\chi$ we have by substitution into equations (2.3) and (2.8) that $y = +\infty$ and $t + f(r_o) = +\infty$. This is proper for a M_1 solution. Expanding equations (2.3) and (2.8) in the neighborhood of $\xi = +\chi$ we find to the

first terms that $y \approx \frac{K(r_o)}{\rho'(\chi)} \frac{1}{(\xi - \chi)}$ and that $t + f(r_o) \approx \frac{1}{\rho'(\chi)} \log (\xi - \chi)$. Since for the particular χ chosen, we know from equation (2.2) that $\rho'(\chi) = -(\Lambda/3)^{\frac{1}{2}}$, we can combine the two parametric expressions into the explicit statement that $y \approx K\sqrt{\frac{3}{\Lambda}}\, e^{\sqrt{\frac{\Lambda}{3}}(t+f)}$ which will be valid in the neighborhood of $y \to +\infty$ and $t \to +\infty$. If we define a new function of integration as $f^*(r_o) = f(r_o) + \frac{1}{2}(3/\Lambda)^{\frac{1}{2}} \log\left(\frac{\Lambda K^2}{3}\right)$ the limiting behavior at infinity is identical with that given in the preliminary discussion in the early part of this chapter. At the other pole $\xi = -\chi$, we have by substitution that $y = +\infty$ and $t + f(r_o) = -\infty$. The limiting behavior of $y(r_o, t \to -\infty)$ can be found by the same procedure and is again identical with that given in the preliminary discussion. Moreover since equation (2.3) can be infinite only at $\xi = \pm\chi$ and congruent points, this behavior as $y \to +\infty$ must be true for all valid M_1 and M_2 solutions.

Since we have taken $\Lambda > 0$, the condition $\alpha^3 > -\Lambda/48$ is true for all positive values of $\alpha(r_o)$. The sign of $\alpha(r_o)$ is the same as the sign of $g(r_o)$; hence the M_1 solution is valid for $\Lambda > 0$, $g(r_o) > 0$, and $h(r_o) > 0$ which are the conditions for solution A in Table I. The condition $\alpha^3 > -\Lambda/48$ allows $\alpha(r_o)$ to assume a limited range of negative values.

But by the definition of $\alpha(r_0)$ the condition $\alpha^3 > -\frac{1}{48}\Lambda$ is identical with the earlier condition $|g^3| < \frac{9}{4}\Lambda h^2$ so with $\Lambda > 0$, $g(r_0) < 0$, and $h(r_0) > 0$ under the condition $|g^3| < \frac{9}{4}\Lambda h^2$ we have solution D of Table I.

Now consider the second class of solutions when $2\omega_1 - \chi \geqslant \xi \geqslant +\chi$ and all congruent segments. Here $\wp(\xi) \leqslant \alpha(r_0)$ and so $K(r_0)$ and hence $h(r_0)$ must be negative to give physically plausible values for y in equation (2.3). We may, if we wish, adjust $F(\xi_0)$ as discussed earlier so that $t + f(r_0) = 0$ at $\xi = \omega_1$. At this point y has its minimum value of $y = K(r_0)\,(e_1 - \alpha)^{-1}$. At the point $\xi = +\chi$, we have $y = +\infty$ and $t + f(r_0) = +\infty$. At the point $\xi = 2\omega_1 - \chi$, we have $y = +\infty$ and $t + f(r_0) = -\infty$. Thus for this range of the parameter, the solution is indeed M_2. The limiting behavior of y when $y \longrightarrow +\infty$ is the same as that discussed in a previous paragraph. Hence we have an M_2 solution for $\Lambda > 0$, $g(r_0) > 0$, and $h(r_0) < 0$ which is exactly solution E in Table I. We also have an M_2 solution for $\Lambda > 0$, $g(r_0) < 0$, $h(r_0) < 0$ under the condition that $|g^3| < \frac{9}{4}\Lambda h^2$, which is part of solution F in Table I.

Let us now consider the second case in Table II where $\Lambda > 0$ and where $\alpha^3 < -\Lambda/48$. Here we have three real roots for the cubic of equation (2.2) and hence will have the allowed system of lines for the variation of ξ

shown in Fig. 2.2. Since $e_2 \leqslant \alpha \leqslant e_3$, χ and its congruent points will again lie on the allowed lines for ξ. To be specific, let us take $\omega_2 \leqslant \chi \leqslant \omega_1 + \omega_2$ and then the situation in the ξ-plane is shown in Fig. 2.4 where again only part of one period-parallelogram is shown. Here there are three classes of solutions according to whether ξ lies within one of three different domains. If ξ ranges over all of the real values or over a congruent domain, O_1 solutions are produced. Since everywhere within these congruent domains $\wp(\xi) > \alpha(r_o)$, $K(r_o)$ and hence $h(r_o)$ will be positive in equation (2.3). At $\xi = 0$ and at all congruent points $y = 0$ by substitution into equation (2.3). If $F(\xi_o) = 0$, then $t + f(r_o) = 0$ at $\xi = 0$. At the congruent point $\xi = 2\omega_1$, $t + f(r_o) = T$ where $T = \frac{4}{\wp'(\chi)}\left[\omega_1 \zeta(\chi) - \chi\, \zeta(\omega_1)\right]$. Let us consider the parametric point ξ within the first period-parallelogram, that is $0 < \xi < 2\omega_1$, and a congruent point $\xi + 2n\omega_1$ where n is any positive or negative integer. Then, upon substitution into equations (2.3) and (2.8) we have $y(r_o, \xi) = y(r_o, \xi + 2n\omega_1)$ and $t(r_o, \xi + 2n\omega_1) = t(r_o, \xi) + nT$. Thus the O_1 solutions are indeed cyclic with a period of T in time. In these solutions the maximum value of $y = K(r_o)\,(e_1 - \alpha)^{-1}$ occurs at $\xi = \omega_1$ and at all congruent points. Since the condition $\alpha^3 < -\Lambda/48$ is

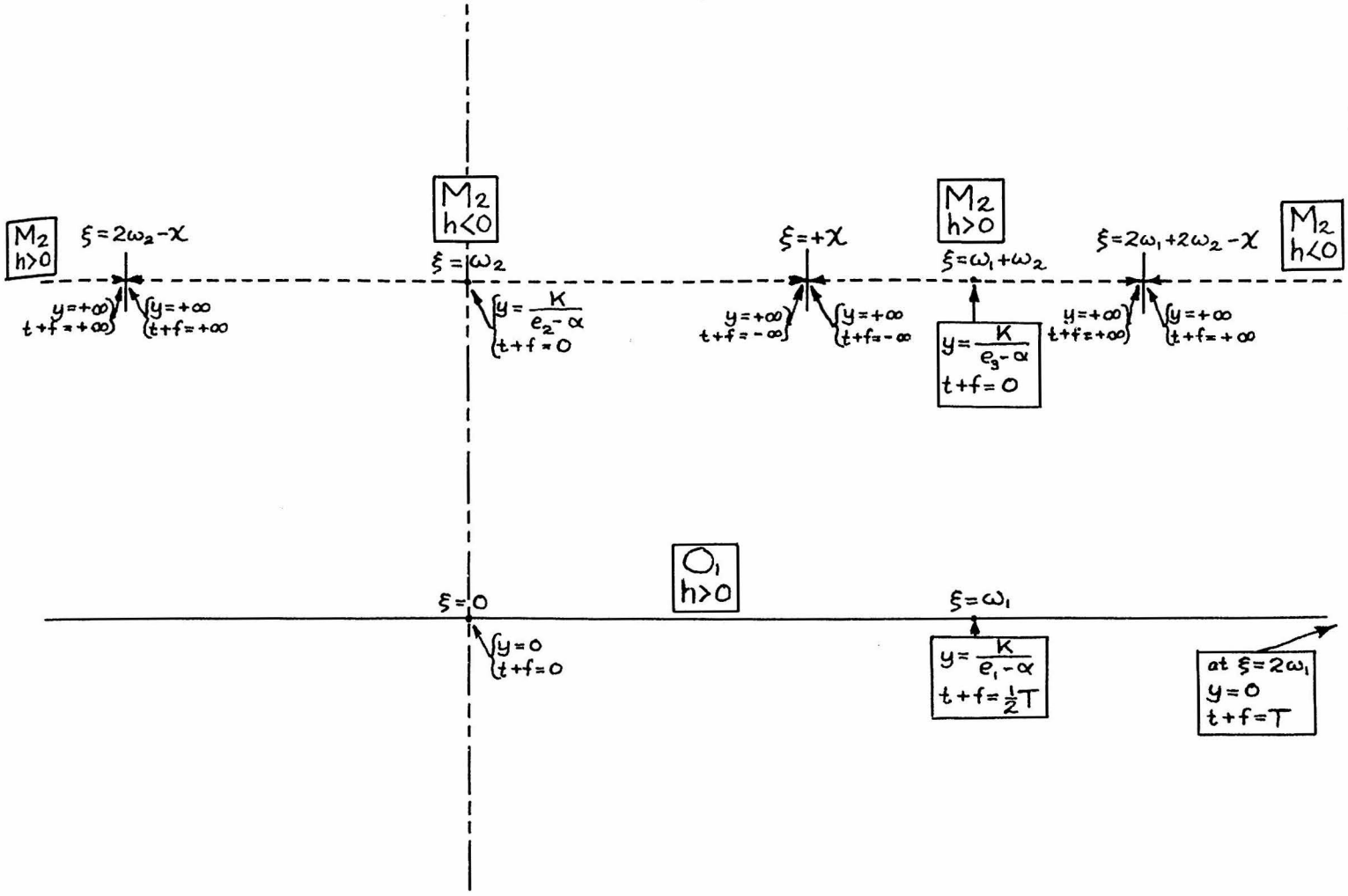

Fig. 2.4

equivalent to the earlier $|g^3| > \frac{9}{4} \Lambda h^2$, these O_1 solutions exist when $\Lambda > 0$, $g(r_0) < 0$, $h(r_0) > 0$ and subject to the stated condition. This is exactly solution B in Table I.

If ξ lies within the domain $2\omega_1 + 2\omega_2 - \chi \geqslant \xi \geqslant +\chi$, or some congruent domain, then M_2 solutions exist. Within

any of these congruent domains $\wp(\xi) \geqslant \alpha(r_o)$, and hence $h(r_o)$ must be positive. At $\xi = 2\omega_1 + 2\omega_2 - \chi$ we find that $y = +\infty$ and $t + f(r_o) = +\infty$. While at $\xi = +\chi$ we have $y = +\infty$ and $t + f(r_o) = -\infty$. At $\xi = \omega_1 + \omega_2$ we have the minimum value of $y = K(r_o)\,(e_3 - \alpha)^{-1}$. We may adjust $F(\xi_o)$ as indicated earlier so that $t + f(r_o) = 0$ at this point, if desired. Therefore the local behavior of these solutions is M_2 as stated. These solutions exist when $\Lambda > 0$, $g(r_o) < 0$, $h(r_o) > 0$, under the condition that $|g^3| > \frac{9}{4}\Lambda h^2$ and are accordingly identical with solution C in Table I.

The remaining class of solutions is produced when ξ lies within the domain $+\chi \geqslant \xi \geqslant 2\omega_2 - \chi$ or any congruent domain. These solutions are also M_2. Within any of these congruent domains $\wp(\xi) \leqslant \alpha(r_o)$, and hence $h(r_o)$ must be negative. The behavior at the poles of y and at the minimum values of y is similar to that of other M_2 solutions. There is no particular reason for repeating this discussion with only small changes in the symbols used. It It might be noted, however, for all solutions in which $h(r_o) < 0$ that the parameter ξ must progress in a negative direction if the time t is to progress in a positive direction. This last class of M_2 solutions exists when $\Lambda > 0$, $g(r_o) < 0$, $h(r_o) < 0$, under the condition $|g^3| > \frac{9}{4}\Lambda h^2$. Hence this is the remainder of the solution F

given in Table I.

In the third case in Table II where $\Lambda < 0$ and $\alpha^3 > -\Lambda/48$, the cubic has three real roots and hence the system of allowed lines for ξ will be identical with those of the previous case. Since $e_3 \leqslant \alpha \leqslant e_1$, χ and its congruent points will not lie on any of the allowed lines. Hence the value of y in equation (2.3) will not possess poles for any of the allowed variations of ξ. Consequently our solutions will be bounded from above in accordance with our early discussion for all cases in which $\Lambda < 0$. Since χ does not lie upon any of the allowed lines for ξ, its exact definition is not important for the determination of the domains of ξ for the various possible classes of solution. For convenience in our drawing, we shall take χ as lying within the fundamental period-parallelogram, but this assumption has no further implications. The situation for this case within the ξ-plane is shown in Fig. 2.5. There are two domains for ξ which will give two different types of solutions. When ξ ranges through all the real values and for all other congruent domains, an O_1 solution exists. When the parameter ξ ranges along a line through ω_2 parallel to the real axis and for all congruent domains, an O_2 solution exists.

For real values of ξ or for any congruent values $\wp(\xi) > \alpha(r_o)$, and hence $h(r_o)$ is positive for these

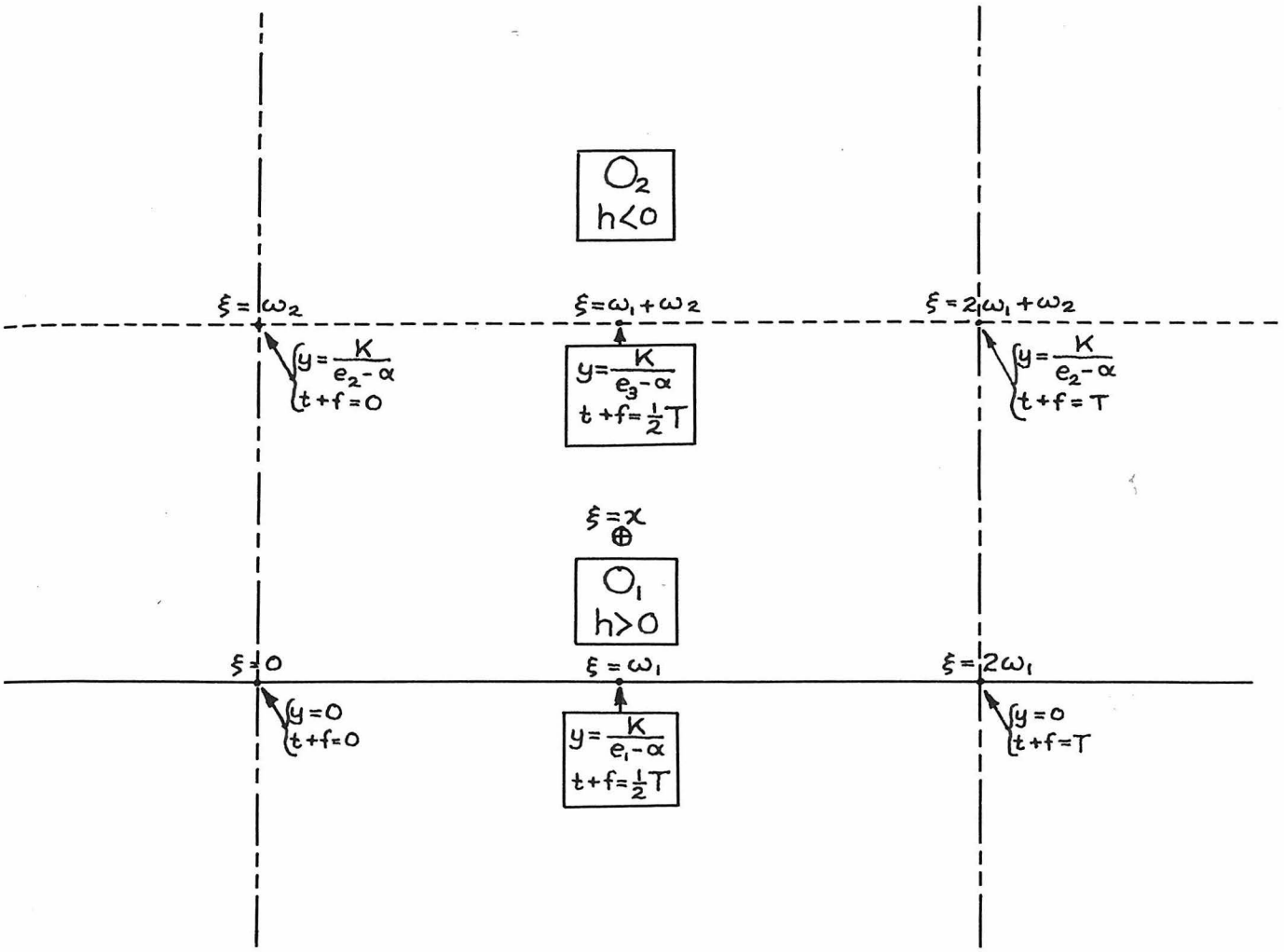

Fig. 2.5

classes. The behavior of these O_1 solutions is identical with that of the O_1 solution discussed earlier, and so it does not need to be repeated here. The condition $\alpha^3 > -\Lambda/48$ is identical with the earlier condition $g^3 > \frac{9}{4}|\Lambda| h^2$. Therefore the O_1 solutions occur when $\Lambda < 0$, $g(r_0) > 0$, $h(r_0) > 0$, subject to the condition $g^3 > \frac{9}{4}|\Lambda| h^2$ and are thereby identified with part of the solution G in Table I.

When ξ lies within the other domain $\wp(\xi) < \alpha(r_o)$, and hence $h(r_o)$ must be negative to give proper physical meaning. Here we have a new class of solutions. Here the minimum value $y = K(r_o)\,(e_2 - \alpha)^{-1}$ occurs at $\xi = \omega_2$ and all congruent points, while the maximum value $y = K(r_o)\,(e_3 - \alpha)^{-1}$ occurs at $\xi = \omega_1 + \omega_2$ and all congruent points. By adjusting the constant $F(\xi_o)$ we can have $t + f(r_o) = 0$ at any desired point, wherein it repeats itself except for an additive constant T in each successive mesh of the period-parallelograms. As in the O_1 solution, we find for a given point ξ and one of its congruent points $\xi + 2n\omega_1$ that $y(r_o, \xi + 2n\omega_1) = y(r_o, \xi)$ and that $t(r_o, \xi + 2n\omega_1) = t(r_o, \xi) - nT$. Hence these solutions are bounded at finite non-zero values both for maximum and minimum values and are periodic with the period T in time. Thus these solutions behave locally like O_2. These solutions occur when $\Lambda < 0$, $g(r_o) > 0$, $h(r_o) < 0$, and subject to the condition $g^3 > \frac{9}{4}|\Lambda| h^2$ and hence are identical with solution I in Table I.

The remaining case in Table II is for $\Lambda < 0$ with $\alpha^3 < -\Lambda/48$. Here we have one real and two complex roots to the cubic. The system of allowed lines for variation of ξ is accordingly that of Fig. 2.1. Since $\alpha < e_1$, χ and its congruent points do not lie within the

allowed range of variation of ξ. Taking χ for convenience in illustration as being within the fundamental period-parallelogram, the situation in the ξ-plane for this last case is that shown in Fig. 2.6. Here there is only one congruent group of domains for which 0_1 solutions exist. Since $\wp(\xi) > \alpha(r_o)$, $h(r_o)$ is positive in value. The condition $\alpha^3 < -\Lambda/48$ is satisfied for all negative values of $\alpha(r_o)$ and consequently for all negative values of $g(r_o)$. Therefore we will expect an 0_1 solution when $\Lambda < 0$, $g(r_o) < 0$, $h(r_o) > 0$ which are the conditions for the solution H of Table I.

The condition $\alpha^3 < -\Lambda/48$ allows $\alpha(r_o)$ to assume a limited range of positive values. This condition is equivalent to the earlier condition $g^3 < \frac{9}{4}|\Lambda| h^2$. Hence with $\Lambda < 0$, $g(r_o) > 0$, $h(r_o) > 0$ and subject to this last condition, we have the remainder of solution G in Table I.

The foregoing comprise all of the physically valid solutions to the original differential equation (2.1) under the most general conditions. However, there are a number of special solutions to this differential equation in terms of non-elliptic functions. These special solutions, which exist under restricted conditions, will be considered in

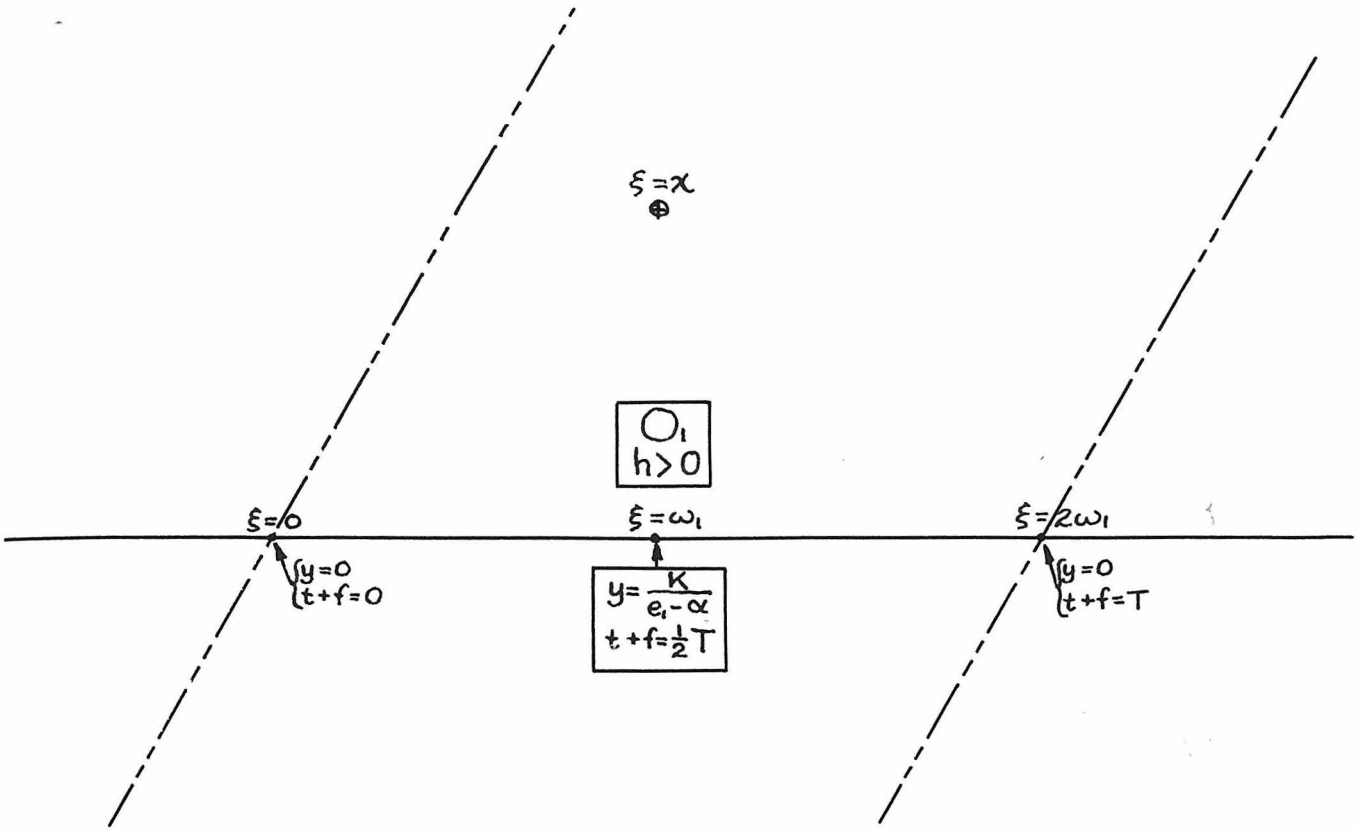

Fig. 2.6

the following chapters. The general solutions agree with the deductions made earlier from the equation (2.1). The only solutions which include the value $y = 0$ are the O_1 and M_1 solutions for which in all cases $h(r_o) > 0$. The only general solutions in which y approaches infinity in value are the M_1 and M_2 solutions for which in all cases $\Lambda > 0$. Moreover, these solutions behave in the neighborhoods of $y = 0$ and of $y \longrightarrow +\infty$ as predicted. When the nomenclature is transformed back to the original

Table III

Λ	Conditions	$g(r_o)$	$h(r_o)$	Type	Parameter Range	Code Letter
+	$\alpha^3 > -\frac{1}{48}\Lambda$	+	+	M_1	$-\chi \leqslant \xi \leqslant +\chi$	A
		+	−	M_2	$2\omega_1 - \chi \geqslant \xi \geqslant +\chi$	E
		−	+	M_1	$-\chi \leqslant \xi \leqslant +\chi$	D
		−	−	M_2	$2\omega_1 - \chi \geqslant \xi \geqslant +\chi$	Part of F
	$\alpha^3 < -\frac{1}{48}\Lambda$	−	+	O_1	ξ = all Real nrs.	B
		−	+	M_2	$\chi \leqslant \xi \leqslant 2\omega_1 + 2\omega_2 - \chi$	C
		−	−	M_2	$\chi \geqslant \xi \geqslant 2\omega_2 - \chi$	Part of F
−	$\alpha^3 > -\frac{1}{48}\Lambda$	+	+	O_1	ξ = all Real nrs.	Part of G
		+	−	O_2	$\xi = \omega_2 +$ Real nrs.	I
	$\alpha^3 < -\frac{1}{48}\Lambda$	+	+	O_1	ξ = all Real nrs.	Part of G
		−	+	O_1	ξ = all Real nrs.	H

functions, the local behavior of the general solutions is found to be that postulated in Table I for the various signs and conditions on Λ, $g(r_o)$, and $h(r_o)$. We have in addition established the actual ranges for the parameter ξ under these conditions so that actual calculations, albeit with great difficulty, can be made.

This material can be presented more succinctly in tabular form, as in Table III. Here the first few columns give the various possible conditions for Λ, $g(r_o)$, and $h(r_o)$. The next column gives the

Robertson-Tolman symbol for the local behavior of the solution in the neighborhood of $r = r_0$. The following column gives one possible range for the parameter ξ for these conditions. The parameter range given is only one out of an infinite number of possible parameter ranges which will fit the same condition. If ξ varies over one of the ranges given in Table III, then $\xi + 2n\omega_1 + 2m\omega_2$ is another congruent parameter range which will give equivalent mathematical results. The last column carries the code letters first assigned to the tenative solutions in Table I. These code letters facilitate the intercomparison of Tables I and III.

Chapter III

Approximate Solutions

In the preceding chapter we have solved the partial differential equation (2.1) in the form of two parametric equations involving Weierstrassian elliptic functions. The allowable values of the parameter which will yield all physically acceptable solutions under the most general conditions have been tabulated in Table III. While these solutions are exact, they are not convenient for actual calculations because of the complexities of the Weierstrassian functions. For the purposes of a numerical computation we would prefer a good non-elliptic approximation to an exact, but intractable, elliptic solution. In this chapter we shall develop some non-elliptic approximations to the general solution under certain conditions upon the values of $\alpha(r)$ and Λ which may or may not be true.

In the preceding chapter we have taken the cosmological constant Λ as being an unrestricted real constant. That is, $-\infty \leqslant \Lambda \leqslant +\infty$. This is in accordance with the general theory of relativity in which Λ may be assigned any real value and the Einstein Field Equation will still have a zero divergence. However, the actual value of the cosmological constant is to be determined from the architecture of the space-time continuum with which we are concerned. The planetary orbits are given within the

limits of the observational errors by the flat assumption that $\Lambda = 0$. Thus the cosmological constant is very small, if not actually zero. Since the homogeneous cosmology is a first approximation to the non-homogeneous models being considered here, the value of the cosmological constant must fall within the range deduced by Tolman (Ref. 1, p. 474) of $-2 \times 10^{-18} < \Lambda < 5.7 \times 10^{-18}$, where the unit of length used is the light-year.

While a plausible assumption is that $\Lambda = 0$, this is a severe restriction, which will be considered in the following chapter. In this chapter we shall make the less severe assumption that Λ is a vanishingly small number, and in particular that $|\Lambda| < |\alpha^3|$. This last assumption is one that may or may not be true since $\alpha(r)$ itself has small values within the surrounding space that we are able to observe. We shall later see that taking $h(r) = 4k^3r^3$ and $g(r) = -r^2/R^2$, where $k^3 = \frac{2}{3}\pi\rho$, reduces the line-element (1.23) to one of the standard forms for a non-static homogeneous line-element. Hence to the homogeneous approximation, we have from Tolman's data (Ref. 1, pp. 461 & 474) that $-2.7 \times 10^{-13} < \alpha^3(r) < 2.9 \times 10^{-14}$, where again the unit of length used is the light-year. This, of course, does not prove the assumption that $|\Lambda| < |\alpha^3|$, but it does show that the condition is possibly true.

If $\Lambda = 0$, then $g_3 = -8\alpha^3$ in (2.2) and the cubic has two repeated roots as shown in (2.14). Under these

conditions the general solution of (2.3) and (2.8) may be expressed in terms of non-elliptic functions. Hence if Λ were vanishingly small and $|\Lambda| < |\alpha^3|$, we would have $g_3 \approx -8\alpha^3$ and the cubic of (2.2) would have two nearly equal roots. Thus we would expect to obtain an approximation to (2.3) and (2.8) in terms of non-elliptic functions, where the approximation becomes better and better as the two nearly equal roots approach each other in value, that is as $\Lambda \longrightarrow 0$.

From Madelung (Ref. 7, p. 78, eqs. 21 & 22 corrected) we take the relations between the invariants and the periods of the Weierstrassian functions as

$$g_2 = \left(\frac{\pi}{\omega_1^x}\right)^4 \left[\frac{1}{12} + 20 \sum_{n=1}^{\infty} \frac{n^3 q^{2n}}{1 - q^{2n}}\right] \tag{3.1}$$

$$g_3 = \left(\frac{\pi}{\omega_1^x}\right)^6 \left[\frac{1}{216} - \frac{7}{3} \sum_{n=1}^{\infty} \frac{n^5 q^{2n}}{1 - q^{2n}}\right] \tag{3.2}$$

where $q = e^{i\pi \frac{\omega_2^x}{\omega_1^x}}$

The periods ω_1^x and ω_2^x are marked with an asterisk to distinguish them from the periods ω_1 and ω_2 used in the previous chapter. The periods used in Chapter II were defined in a particular way while ω_1^x and ω_2^x may be any one of the six possible pairs of values. We shall identify ω_1^x and ω_2^x to suit our convenience. In some cases they will be defined as in Chapter II while in other cases the indentification will be exactly reversed

from that of the previous chapter.

As Λ approaches zero, q approaches either zero or unity, depending upon how the two periods are defined. That is, as $\Lambda \longrightarrow 0$, one of the periods approaches a finite limit, which may be either real or imaginary, while the other period approaches an infinte limit, which again may be either real or imaginary. We shall take ω_1^x as the period whose limit is finite as $\Lambda \longrightarrow 0$. Then by the definition of q, $q \longrightarrow 0$ as $\Lambda \longrightarrow 0$. Consequently, if we expand the Weierstrassian functions in powers of q, the first few terms of these expansions will give a good approximation for computational purposes as $\Lambda \longrightarrow 0$. In this work we shall carry out the approximation as far as the q^2 term wherever this is feasible.

Since our invariants are $g_2 = 12\alpha^2$ and $g_3 = -8\alpha^3 - \Lambda/3$, we can use equations (3.1) and (3.2) to obtain approximate expressions for $\alpha(r)$, ω_1^x, and ω_2^x. The uncertainity in sign upon taking the square root of (3.1) may be resolved by an appeal to the degenerations of the Weierstrassian functions (vide Ref. 6, p. 105). When $q = 0$ with $\alpha < 0$ we find that ω_1^x is a real number; hence the negative sign must be used and we have

$$\alpha(r) \approx -\frac{1}{3}\left(\frac{\pi}{2\omega_1^x}\right)^2\left[1 + 120\,q^2 + - - -\right] \qquad (3.3)$$

This approximation may in turn be solved for the period ω_1^x

$$\omega_1^x \approx \frac{\pi}{2\sqrt{-3\alpha}}\left[1 + 60\,q^2 + - - -\right] \tag{3.4}$$

While the definition of q gives the approximation for ω_2^x as

$$\omega_2^x \approx \frac{1}{2\sqrt{3\alpha}}\left[1 + 60q^2 + - - -\right] \log q \tag{3.5}$$

An approximation is also needed for q if computations are to be made. This approximation can be made by forming $3\sqrt{3}\ g_3\, g_2^{-3/2} = -1 - \Lambda/24\alpha^3$ from equations (3.1) and (3.2), rearranging, taking a square root, and then inverting the series as

$$q \approx \frac{1}{144}\sqrt{\frac{-\Lambda}{\alpha^3}}\left[1 - 5.295\times10^{-3}\,\frac{\Lambda}{\alpha^3} + - - -\right] \tag{3.6}$$

Therefore, in the extreme condition that $\alpha^3 = -\Lambda$, the first term of (3.6) by itself gives q to two significant figures. If $|\Lambda|$ is appreciably smaller than $|\alpha^3|$, the first term of (3.6) alone gives q to sufficient accuracy for all computational purposes with the present observational data. By taking the condition $\alpha^3 = -\Lambda$ as extreme, then $|q|$ will have a maximum value of 6.94×10^{-3}. Thus approximations to the Weierstrassian functions which are

Table IV

Class	Λ	$\alpha(r)$	ξ = Real	$\xi = \omega_2$ + Real	q	ω_1^*	ω_2^*
I	+	+	M_1 M_2	—	$i\epsilon$	$-iC_1$	$C_2 - \frac{i}{2}C_1$
II		−	O_1	M_2 M_2	ϵ	C_1	iC_2
III	−	+	O_1	O_2	ϵ	$-iC_1$	C_2
IV		−	O_1	—	$i\epsilon$	C_1	$iC_2 + \frac{1}{2}C_1$

carried only to the q^2 term will give more significant figures than the present observational data would justify. Hence the approximations to the general solutions, which are to be derived, will be quite good under the conditions stated at the beginning of this chapter.

The three approximations of (3.4), (3.5), and (3.6) may be used to find nearly exact values of q, ω_1^*, ω_2^* for the four main groups of conditions listed in Table III. As such, the data of Table IV are merely an extension of the data of Table III and are valid only under the conditions of this chapter. In the first column numerals are assigned to the four groups so that they may be referred to in the text. The second column gives the sign of the cosmological

constant. The third column gives the sign of $\alpha(r)$, which is sufficient in the present circumstances where $|\Lambda| < |\alpha^3|$. The next two columns are only reminders as to the types of solutions which will occur. The exact statements about the types of solutions which will exist are found in Table III. The sixth column gives the approximate value of q, where ϵ is a small finite real number whose value is to be determined from (3.6). The following column gives the approximate value of ω_1^x, where C_1 is a finite real number whose value is to be derived from (3.4). The last column gives the approximate value of ω_2^x, where C_2 is a large real number which approaches infinity as q approaches zero. The exact value of C_2 is to be found from (3.5). Table IV is actually only a mnemonic to aid in arranging the various approximations to the general solutions.

There is considerable freedom in specifying the periods ω_1^x and ω_2^x because of the ambiguity in sign of equations (3.4) and (3.5) and because of the symmetry of the poles of $\wp(\xi)$ about the real axis. Those listed in Table IV were chosen to give a vanishing q with an infinitely increasing C_2. It has been assumed here that both C_1 and C_2 are positive real numbers. Equation (3.5) gives complex values for ω_2^x in classes I and IV since $\log i\epsilon = \frac{1}{2}\pi i + \log \epsilon$. The periods of Table IV

are readily identified with the periods of the previous chapter. They are, in fact

I	$\omega_1^x = \omega_1 - 2\omega_2$	$\omega_2^x = \omega_1 - \omega_2$
II	$\omega_1^x = \omega_1$	$\omega_2^x = \omega_2$
III	$\omega_1^x = -\omega_2$	$\omega_2^x = \omega_1$
IV	$\omega_1^x = \omega_1$	$\omega_2^x = \omega_2$

In classes II and III the three roots of the cubic in (2.2) are all real and the period-parallelogram is a highly elongated rectangle which approaches an infinite strip as q approaches zero. In classes I and IV two of the roots are complex and the elongated period-parallelogram is not rectangular but still approaches an infinite strip parallel either to the real or to the imaginary axis as q approaches zero. In class IV it is immediately obvious that Real $\omega_2 = \frac{1}{2}$ Real ω_1 as stated in the previous chapter. If the periods for class I are solved, we find that $\omega_1 = 2C_2$ and $\omega_2 = C_2 + \frac{1}{2}iC_1$. Hence here also Real $\omega_2 = \frac{1}{2}$ Real ω_1. This is a verification for the statement in chapter II, which was derived mainly by geometrical reasoning.

The Weierstrassian Pe-function may be stated as (vide Ref. 5, p. 379)

$$\wp(\xi) = \left(\frac{\pi}{2\omega_1^x}\right)^2 \left\{ -\frac{1}{3} + \sum_{-\infty}^{+\infty} \csc^2 \frac{\pi(\xi - 2n\omega_2^x)}{2\omega_1^x} - \sum_{-\infty}^{+\infty\,\prime} \csc^2 \frac{n\pi\omega_2^x}{\omega_1^x} \right\} \qquad (3.7)$$

where the prime over the second summation sign signifies that the $n = 0$ term is to be suppressed. Equation (3.7) may be used with equation (3.3) to form approximations for equation (2.3). The latter approximation may in turn be integrated according to equation (2.5) to give an approximation for equation (2.8). This we shall now proceed to do for each of the eight allowable solutions listed in Table IV.

1. Classes II & IV, $\xi = x$.

We shall use x throughout this chapter to signify a purely real variable. Here we wish to approximate in the neighborhood of $\xi = 0$ and along the real axis. We shall accordingly find approximations to the following two solutions

II. $\Lambda > 0$ $g < 0$ $h > 0$ $\omega_1^* = C_1$ $q = \epsilon$ $q^2 = +\epsilon^2$ 0_1

IV. $\Lambda < 0$ $g < 0$ $h > 0$ $\omega_1^* = C_1$ $q = i\epsilon$ $q^2 = -\epsilon^2$ 0_1

From (3.7) and (3.3) to the q^2 term, we find

$$\rho(\xi) - \alpha \approx \left(\frac{\pi}{2C_1}\right)^2 \left\{ \frac{1}{\sin^2 \frac{\pi x}{2C_1}} + 8q^2\left(6 - \cos\frac{\pi x}{C_1}\right) + - - - \right\} \qquad (3.8)$$

Under the present restrictions, $56\, q^2 \ll 1$, so (3.8) may be used in (2.3) to give an approximate solution which is valid over the entire real range of x. Carrying out the approximations and transforming the coefficients back into the original $g(r)$ and $h(r)$ functions, we have

$$y \approx \frac{h}{2(-g)}\left[1 - \cos\Theta + q^2\{80 - 65\cos\Theta - 16\cos 2\Theta + \cos 3\Theta\} + - - -\right] \qquad (3.9)$$

$$t + f \approx \frac{h}{2(-g)^{3/2}}\left[\Theta - \sin\Theta + q^2\{140\,\Theta - 125\sin\Theta - 8\sin 2\Theta + \tfrac{1}{3}\sin 3\Theta\} + - - -\right] \qquad (3.10)$$

where $\Theta = \frac{\pi x}{C_1}$

with $q^2 = +\epsilon^2$, we have Solution B of Tables I & III;

with $q^2 = -\epsilon^2$, we have Solution H of Tables I & III.

The relation between y $(-g)^{\frac{1}{2}}(t + f)$ is a distorted cycloid, approaching more and more closely to a perfect cycloid as $q \longrightarrow 0$. This is in agreement with de Sitter's (Ref. 8) numerical quadratures of the non-static homogeneous models.

2. Classes I & III, $\xi = x$.

We wish to find the approximations in the neighborhood of $\xi = 0$ and along the real axis for the following solutions

I. $\Lambda > 0$ $g > 0$ $h > 0$ $\omega_1^* = -iC_1$ $q = i\epsilon$ $q^2 = -\epsilon^2$ M_1

III. $\Lambda < 0$ $g > 0$ $h > 0$ $\omega_1^* = -iC_1$ $q = \epsilon$ $q^2 = +\epsilon^2$ Q_1

Carrying the series expansion to the q^2 term, we find that

$$\wp(\xi) - \alpha \approx -\left(\frac{\pi}{2C_1}\right)^2\left\{-\frac{1}{\sinh^2\frac{\pi x}{2C_1}} + 8q^2\left(6 - \cosh\frac{\pi x}{C_1}\right) + - - -\right\} \qquad (3.11)$$

This cannot be used in (2.3) to give an approximation over the entire range of x since q^2 term will exceed the first term in value as $x \rightarrow \pm\infty$. We shall first approximate in the vicinity of $x = 0$ where the q^2 term is appreciably smaller than the first term. Then we must consider a new approximation in the neighborhood of $\xi = \omega_1$ for the further

behavior of the solutions. We note, however, that for class III $q^2 = +\epsilon^2$ so (3.11) does not change sign as x varies throughout its real range. Hence both approximations for this class are for a single cyclic solution. In class I $q^2 = -\epsilon^2$ so (3.11) will change sign, going through a zero, which produces a pole for $y(r,t)$ at $x = \chi$ and at any congruent point. Thus the two approximations are for two different solutions. The approximation in the neighborhood of $x = 0$ is for the M_1 solution with $h > 0$, while the other approximation in the vicinity of $\xi = \omega_1$ will yield the M_2 solution with $h < 0$. Carrying out the indicated approximations in the neighborhood of $x = 0$, we have

$$y \approx \frac{h}{2g}\left[\cosh\theta - 1 - q^2\{80 - 65\cosh\theta - 16\cosh 2\theta + \cosh 3\theta\} + - - -\right] \qquad (3.12)$$

$$t+f \approx \frac{h}{2g^{3/2}}\left[\sinh\theta - \theta - q^2\{140 - 125\sinh\theta - 8\sinh 2\theta + \tfrac{1}{3}\sinh 3\theta\} + - - -\right] \qquad (3.13)$$

where $\theta = \dfrac{\pi x}{C_1}$

with $q^2 = +\epsilon^2$, we have part of Solution G of Tables I & III;

with $q^2 = -\epsilon^2$, we have Solution A of Tables I & III.

If these parametric equations are plotted with $q = 0$, the resulting curve for the relation between y and $t + f$ will be found to be similar to de Sitter's Curve V as given by Tolman (Ref. 1, p. 411). This should be expected, of course, since de Sitter's curve V is a numerical integration of a M_1

solution in a homogeneous model. These parametric equations might be termed a distorted imaginary cycloid, since the substitution of an imaginary parameter and a change of sign for $g(r)$ into (3.12) and (3.13) will exactly yield the earlier parametric equations (3.9) and (3.10).

For the O_1 solution of class III, $\omega_1 = C_2 \longrightarrow +\infty$ as $q \longrightarrow 0$. Consequently, we would expect this O_1 solution to resemble more and more closely a M_1 solution as $q \longrightarrow 0$. For the O_1 approximation in the neighborhood of $\xi = \omega_1$ which will be presented shortly, we expect to find less and less physical meaning as $q \longrightarrow 0$, degenerating into $y = +\infty$ and $t + f = \pm\infty$ when $q = 0$. Thus, as the limit is approached, the parametric approximations of (3.12) and (3.13) will be sufficient for the entire solution of class III.

3. Class III, $\xi = \omega_1 + x$.

In the previous section we have seen that the behavior of the O_1 solution in the neighborhood of $\xi = 0$ and congruent points is very similar to the behavior of a M_1 solution. We need to approximate the O_1 solution in the neighborhood of its maximum points, $\xi = \omega_1$ and congruent points, to show that it is actually cyclic. Since $\omega_1 = \omega_2^x$ in class III, we shall approximate up to the q^2 term in the vicinity of $\xi = \omega_2^x + x$, which gives

$$\wp(\xi) - \alpha \approx -\left(\frac{\pi}{2C_1}\right)^2 \left\{ -8q \cosh \frac{\pi x}{C_1} + 32q^2 \left(2 - \cosh^2 \frac{\pi x}{C_1}\right) + - - - \right\} \qquad (3.14)$$

Using (3.14) in turn, we find for the O_1 solution in the neighborhood of $\xi = \omega_1 + x$ that

$$y \approx \frac{h}{g}\left[\frac{1}{8q} \operatorname{sech} \theta + \operatorname{sech}^2 \theta - \frac{1}{2} + - - -\right] \qquad (3.15)$$

$$t + f \approx \frac{h}{g^{3/2}}\left[\frac{1}{8q} \tan^{-1}(\sinh \theta) + \tanh \theta - \frac{1}{2}\theta + - - -\right] \qquad (3.16)$$

where $\theta = \frac{\pi x}{C_1}$, $q = \epsilon$ and all other condition on the class III O_1 solution are as in the previous section.

As predicted, we see as $q \longrightarrow 0$ that $y \longrightarrow +\infty$ and $t + f \longrightarrow \pm \infty$. We can gain further insight into the behavior of this approximation by assuming that q is sufficiently small so that only the first term in (3.15) and (3.16) needs to be considered. Then the two parametric equations can be solved explicity as

$$y \approx \sqrt{\frac{3q}{-\Lambda}} \cos\left[\sqrt{\frac{-\Lambda}{3}}\,(t + f)\right] \qquad (3.17)$$

Thus the maxima of the O_1 solutions are distorted cosine curves. Hence the O_1 solutions are cyclic as indicated. But it is also obvious when $\Lambda \longrightarrow 0$ that the O_1 solutions approach to M_1 behavior.

4. Class I, $\xi = \omega_1 + x$.

Here we seek the approximation in the neighborhood of $\xi = \omega_1$ and congruent points and along the real axis for the following solution

I. $\Lambda > 0 \quad g > 0 \quad h < 0 \quad \omega_1^* = -iC_1 \quad q = i\epsilon \quad q^2 = -\epsilon^2 M_2$

We cannot use the approximations of the preceding section since in class I we have $\omega_1 = 2\omega_2^x - \omega_1^x$. Carrying out the required approximation up to the q^2 term for the neighborhood of $\xi = 2\omega_2^x - \omega_1^x + x$, we find

$$\wp(\xi) - \alpha \approx -\left(\frac{\pi}{2C_1}\right)^2 \left\{ \frac{1}{\cosh^2 \frac{\pi x}{2C_1}} + 8q^2\left(6 + \cosh\frac{\pi x}{C_1}\right) + - - - \right\} \qquad (3.18)$$

Using this to find the corresponding approximations to the general solution, we have

$$y \approx \frac{-h}{2g}\left[\cosh\theta + 1 + q^2\{60 + 30\cosh\theta - 36\cosh 2\theta - 6\cosh 3\theta\} + - - -\right] \qquad (3.19)$$

$$t + f \approx \frac{-h}{2g^{3/2}}\left[\sinh\theta + \theta + q^2\{120\,\theta + 90\sinh\theta - 18\sinh 2\theta - 2\sinh 3\theta\} + - - -\right] \qquad (3.20)$$

where $\theta = \frac{\pi x}{C_1}$

Under the conditions stated at the beginning of this section, this is Solution E of Tables I & III

The parabola-like relationship between y and t defined by the approximations of (3.19) and (3.20) is not one of the named family of parametric curves. If this relationship is calculated, it will be found to resemble

closely de Sitter's M_2 solutions for the homogenous models, as might be expected. This approximation is not good for large values of Θ, and consequently, for large values of y or $t + f$, since the q^2 term of (3.18) will then exceed the first term in value.

5. Class III, $\xi = \omega_2 + x$.

We have produced the approximations for all of the solutions listed in Table IV in which the parameter ξ ranges over the real axis or some congruent line. There remain three solutions in Table IV in which the parameter varies over $\omega_2 + x$ or some congruent line. Here we wish to consider the O_2 solution which exists under the following conditions:

III. $\Lambda < 0 \quad g > 0 \quad h < 0 \quad \omega_1^* = -iC_1 \quad q = \epsilon \quad q^2 = +\epsilon^2 \quad O_2$

Since in class III $\omega_2 = -\omega_1^x$, we shall approximate up to the q^2 term in the neighborhood of $\xi = \omega_1^x + x$

$$\rho(\xi) - \alpha \approx -\left(\frac{\pi}{2C_1}\right)^2 \left\{ \frac{1}{\cosh^2 \frac{\pi x}{2C_1}} + 8q^2\left(6 + \cosh\frac{\pi x}{C_1}\right) + - - - \right\} \qquad (3.21)$$

This approximation may be used in (2.3) only for a limited range of x since the q^2 term will exceed the first term as $x \longrightarrow \pm\infty$. To obtain the temporal behavior of the entire O_2 solution, we must make a second approximation in the neighborhood of $\xi = \omega_1 + \omega_2 + x$. For the present neighborhood, we have

$$y \approx \frac{-h}{2g}\left[\cosh\theta + 1 + q^2\{60 + 30\cosh\theta - 36\cosh 2\theta - 6\cosh 3\theta\} + - - -\right] \qquad (3.22)$$

$$t + f \approx \frac{-h}{2g^{3/2}}\left[\sinh\theta + \theta + q^2\{120\,\theta + 90\sinh\theta - 18\sinh 2\theta - 2\sinh 3\theta\} + -\right] \qquad (3.23)$$

where $\theta = \frac{\pi x}{C_1}$

This is formally identical with an earlier approximation, except that here $q^2 = +\epsilon^2$ and this part of Solution I of Tables I & III. Consequently, the O_2 solution in the neighborhood of $\xi = \omega_2$ and congruent points resembles the M_2 solution, becoming identical with it when $q = 0$.

6. Class III, $\xi = \omega_1 + \omega_2 + x$.

To continue the study of the O_2 solution, we now want to approximate to it in the vicinity of its maxima at $\xi = \omega_1 + \omega_2$ and congruent points. Carrying out this approximation up to the q^2 term in the neighborhood of $\xi = \omega_1^x + \omega_2^x + x$, we find

$$\wp(\xi) - \alpha \approx -\left(\frac{\pi}{2C_1}\right)^2\left\{8q^2\cosh\frac{\pi x}{C_1} + 32q^2\left(2 - \cosh^2\frac{\pi x}{C_1}\right) + - - -\right\} \qquad (3.24)$$

Using (3.24) gives

$$y \approx \frac{-h}{g}\left[\frac{1}{8q}\operatorname{sech}\theta - \operatorname{sech}^2\theta + \tfrac{1}{2} + - - -\right] \qquad (3.25)$$

$$t + f \approx \frac{-h}{g^{3/2}}\left[\frac{1}{8q}\tan^{-1}(\sinh\theta) - \tanh\theta + \tfrac{1}{2}\theta + - - -\right] \qquad (3.26)$$

where $\theta = \frac{\pi x}{C_1}$

Under the same conditions as the previous section, this is the remainder of Solution I of Tables I & III. By taking q sufficiently small so that all terms other than the first may be ignored, we can solve these parametric equations explicitly as

$$y \approx \sqrt{\frac{3g}{-\Lambda}} \cos\left[\sqrt{\frac{-\Lambda}{3}}\,(t+f)\right] \tag{3.27}$$

Thus the behavior of the O_2 solution at its maxima is similiar to the behavior of the O_1 solution at its maxima, being a distorted cosine curve.

7. Class II, $\xi = \omega_2 + x$.

There are two solutions that exist along this line, both being M_2. We shall first consider the M_2 solution that exists in the neighborhood of $\xi = \omega_2$ under the following conditions:

II. $\Lambda > 0$ $\quad g < 0$ $\quad h < 0$ $\quad \omega_1^* = C_1$ $\quad q = \epsilon$ $\quad q^2 = +\epsilon^2$ $\quad M_2$

In the vicinity of $\xi = \omega_2^x + x$ to the q^2 term, we have

$$\rho(\xi) - \alpha \approx \left(\frac{\pi}{2C_1}\right)^2 \left\{-8q\cos\frac{\pi x}{C_1} + 32q^2\left(2 - \cos^2\frac{\pi x}{C_1}\right) + \cdots\right\} \tag{3.28}$$

Employing (3.28) in the usual way, we find

$$y \approx \frac{(-h)}{(-g)}\left[\frac{1}{8q}\sec 2\theta + \sec^2 2\theta - \frac{1}{2} + \cdots\right] \tag{3.29}$$

$$t+f \approx \frac{(-h)}{(-g)^{3/2}}\left[\frac{1}{8q}\log\tan\left(\theta+\frac{\pi}{4}\right)+\tan 2\theta-\theta+---\right] \tag{3.30}$$

where $\theta = \frac{\pi x}{2C_1}$

Under the conditions stated here, this is part of Solution F in Tables I & III. With q sufficiently small so that all terms other than the first may be dropped, these parametric equations can be solved explicitly by means of the elementary lambda and gudermannian functions as

$$y \approx \sqrt{\frac{-3g}{\Lambda}}\cosh\left[\sqrt{\frac{\Lambda}{3}}(t+f)\right] \tag{3.31}$$

Thus this M_2 solution is a distorted catenary.

8. Class II, $\xi = \omega_1 + \omega_2 + x$.

The remaining solution listed in Table IV is

II. $\Lambda > 0$ $g < 0$ $h > 0$ $\omega_1^* = C_1$ $q = \epsilon$ $q^2 = +\epsilon^2$ M_2.

Carrying out this approximation to the q^2 term in the neighborhood of $\xi = \omega_1^* + \omega_2^* + x$, we find

$$y \approx \frac{h}{(-g)}\left[\frac{1}{8q}\sec 2\theta - \sec^2 2\theta + \frac{1}{2} + ---\right] \tag{3.32}$$

$$t+f \approx \frac{h}{(-g)^{3/2}}\left[\frac{1}{8q}\log\tan\left(\theta+\frac{\pi}{4}\right) - \tan 2\theta + \theta + ---\right] \tag{3.33}$$

where $\theta = \frac{\pi x}{2C_1}$

Under the conditions stated here, this is Solution C of Tables I & III. The further approximation when only the q^{-1} term is considered is identical with (3.31) so this M_2 solution is also a distorted catenary.

The foregoing comprise the approximations to every solution listed in Table IV under the conditions stated at the beginning of this chapter. Approximations have not been given for every solution in Table III. Solution D and parts of Solutions F and G exist only when $|\Lambda| \geqslant 48\,|\alpha^3|$, which is contrary to our approximating conditions. These solutions, as well as those of Table IV, may be approximated under different conditions. Since the cubic of (2.2) will also have repeated roots when $\Lambda = -48\,\alpha^3$, we shall take $\Lambda = -48\alpha^3 + \lambda$, where λ is a vanishingly small real quantity.

With this new condition, we can approximate to the q^2 term, as before. Equation (3.1) gives an approximation for $\alpha(r)$ which is

$$\alpha(r) \approx +\frac{1}{3}\left(\frac{\pi}{2\omega_1^*}\right)^2\left[1 + 120\,q^2 + - - -\right] \tag{3.34}$$

This is identical with (3.3) except for sign. When the square root of (3.1) is taken, the sign is ambigious. As before, the correct sign to use is determined from the degenerations of the elliptic function (vide Ref. 6, p. 105) when $q = 0$. Here, one of the periods is real and finite when $\Lambda < 0$ and $\alpha > 0$. Thus the positive sign is used in this case.

The approximation for $\alpha(r)$ and the definition of q may be used to express approximations for the two

periods. Equation (3.34) gives

$$\omega_1^x \approx \frac{\pi}{2\sqrt{3\alpha}}\left[1 + 60\,q^2 + - - -\right] \tag{3.35}$$

This is identical with (3.4) except for the sign under the radical. Here, as before, we shall take ω_1^x to be that period which approaches a finite limit as $\lambda \longrightarrow 0$. Using (3.35) and the definition of q we find

$$\omega_2^x \approx \frac{1}{2\sqrt{-3\alpha}}\left[1 + 60\,q^2 + - - -\right] \log q \tag{3.36}$$

This is again identical with (3.5) except for the sign under the radical.

The approximation for q is made by forming $3\sqrt{3}\ g_3\, g_2^{-3/2} = 1 - \lambda/24\,\alpha^3$, where the definition of λ has been used. Substituting for g_2 and g_3 from (3.1) and (3.2), rearranging the series, taking a square root of the series and finally inverting the series, we have

$$q \approx \frac{1}{144}\sqrt{\frac{\lambda}{\alpha^3}}\left[1 + 5.295 \times 10^{-3}\,\frac{\lambda}{\alpha^3} + - - -\right] \tag{3.37}$$

which is identical with (3.6) except for some signs and one symbol.

The foregoing approximations along with (3.7) may be used to establish approximations up to the q^2 term for the general solutions (2.3) and (2.8). This approximation

for the basic term of (2.3) is

$$\wp(\xi)-\alpha \approx \left(\frac{\pi}{2\omega_1^x}\right)^2\left[-\frac{2}{3}+\sum_{-\infty}^{+\infty}\csc^2\frac{\pi(\xi-2n\omega_2^x)}{2\omega_1^x} - 32q^2+---\right] \qquad (3.38)$$

As shown in the previous discussion, the remaining cosecant series is to be further expanded in terms of x and q for the particular domain of ξ which is used. The approximation for (2.3) and (2.8) may be carried out in the same way, but not with the same ease, as before. The presence of the $2/3$ term in (3.38) complicates this problem. In the previous approximations the $1/3$ term of (3.7) was exactly cancelled by the same term of opposite sign in (3.3).

The conditions for this second set of approximations have little intuitive appeal since they require a non-vanishing cosmological constant in the most general cases. Furthermore, this appeal is lessened by the difficulty of integrating the approximation in (2.5). Consequently, a group of exact solutions will be given here for the special case in which $q = 0$.

A review of the previous approximations will show that they fall into two classes. Some are made up of a group of terms not containing q and of further terms in q^2 or higher powers. As $q \longrightarrow 0$ these approximations are dominated by the terms that do not contain q. Thus, for a somewhat rougher approximation, as $q \longrightarrow 0$ the terms in q may be disregarded entirely. The remaining class of

approximations has a dominant term involving q^{-1}. Thus as $q \longrightarrow 0$ these approximations have little physical usefulness since the minimum value of y becomes an extremely large number. In the same sense the exact solutions in the special case that $\lambda = 0$ are rough approximations to at least parts of Solutions B, C, D, F, G, and I of Table I when $\Lambda \approx -48\alpha^3$. Taking then, the special case $q = 0$ the approximation of (3.38) reduces to the exact statement of

$$\rho(\xi) - \alpha = \alpha\left[3\csc^2\frac{\pi\xi}{2\omega_1^x} - 2\right] \tag{3.39}$$

Before deriving the special solutions from (3.39) let us turn back to (2.1) and Table I to determine the nature of the expected special solutions. First, let us consider $\Lambda > 0$. Then the special condition $\Lambda = -48\alpha^3$ is identical with $-g^3 = 9\Lambda h^2/4$. With $|g^3| > \frac{9}{4}\Lambda h^2$, as will be true when $q \longrightarrow 0$ from one side, (2.1) has two solutions as shown in the discussion of chapter II. One of these solutions is O_1 with a maximum value of $y = E_1$ and with $\frac{\partial^2 y}{\partial t^2}\Big|_{E_1} < 0$. While the other solution is M_2 with a minimum value of $y = E_2$ and with $\frac{\partial^2 y}{\partial t^2}\Big|_{E_2} > 0$. Then as $q \longrightarrow 0$ we must have $E_1 \longrightarrow E_2$ and $\frac{\partial^2 y}{\partial t^2}\Big|_{E_1} \longrightarrow \frac{\partial^2 y}{\partial t^2}\Big|_{E_2} \longrightarrow 0$. With $q = 0$, $E_1 = E_2 = E''$ where E'' has been used to denote the repeated root of the cubic of (2.1), and

$\left.\frac{\partial y}{\partial t}\right|_{E''} = 0$ along with $\left.\frac{\partial^2 y}{\partial t^2}\right|_{E''} = 0$. An inspection of (2.1) with a repeated root in the cubic will show that not only the first two time derivatives are zero at the repeated root, but also that all higher time derivatives are likewise zero. Consequently, in the limiting solutions to (2.1) y cannot cross the line $y = E''$ nor can y touch the line $y = E''$ as either a maximum or a minimum. Therefore these solutions can only approach or depart from $y = E''$ asymptotically. Thus the O_1 solution, Solution B of Table I, must resemble locally an asymptotic universe of the first type, usually denoted as A_1, as $q \longrightarrow 0$. Likewise the M_2 solution, which is Solution C of Table I, must locally come to resemble closely an asymptotic universe of the second type, which is denoted as A_2.

With $|g^3| < \frac{9}{4} \Lambda h^2$, as will be true when $q \longrightarrow 0$ from the other side, (2.1) has a single M_1 solution, Solution D of Table I. Then as $q \longrightarrow 0$ this M_1 solution must take on the appearance of de Sitter's curve III as given by Tolman (Ref. 1, p. 411) where the quasi-plateau becomes longer and longer as the critical point is approached. At the critical point this M_1 solution must break up into the two special solutions, A_1 and A_2. Thus the A_1 special solution is a rough approximation to Solution D for $0 \leqslant y \leqslant E''$ when $\lambda \longrightarrow 0$. The A_2 special solution is a rough approximation for Solution D with

$E'' \leqslant y \leqslant +\infty$ when $\lambda \longrightarrow 0$. It should be noted that the quasi-plateau of the M_1 solution does not contain a point of inflection since the first time derivative is never zero. This is in agreement with the statement of Chapter II that none of the physically acceptable solutions of (2.1) can have points of inflection, but can only vary monotonically with respect to time between maxima and minima. This is always true since if y is any solution of (2.1), $\frac{\partial^n y}{\partial t^n} = F_1(y)\frac{\partial^2 y}{\partial t^2} + F_2(y)\frac{\partial y}{\partial t}$. Hence for a particular value of y for which $\frac{\partial^2 y}{\partial t^2} = 0$ and $\frac{\partial y}{\partial t} = 0$, y can only approach that value asymptotically since all higher time derivatives are also zero.

Since in this special case at $y = E''$ we have all time derivatives equal to zero, we have a special static solution to (2.1). This static solution is unstable to any perturbations in y. Any perturbation tending to increase y will be accentuated and the local behavior becomes A_2. Any contrary tendency for y to decrease would also be further aggravated into local A_1 behavior. Thus this special static solution behaves locally like the Einstein static homogeneous model and will be denoted as E_1, Einstein behavior of the first type. This extension of the Robertson-Tolman notation is to allow for a later type of static behavior.

Because the sign of $h(r)$ has no effect upon the sign of $\alpha(r)$, the special condition $\Lambda = -48\alpha^3$ can be satisfied with $g < 0$ $h < 0$, which by Table I means an M_2 solution. The cubic of (2.1) has a repeated root under these conditions, but the repeated root is negative and thus has no physical significance. Nevertheless, the cubic of (2.2) has a repeated root and hence (2.3) and (2.8) can be expressed exactly in terms of non-elliptic functions. This gives a non-elliptic statement of Solution F under the specified special conditions.

When $\Lambda < 0$, the special condition $\Lambda = -48\alpha^3$ is identical with $g^3 = -9\Lambda h^2/4$. With $g^3 > -\frac{9}{4}\Lambda h^2$ as will be true when $q \rightarrow 0$ from one side, (2.1) has a single 0_2 solution, Solution I of Table I. This cyclic solution has maxima and minima of E_2 and E_1. As $q \rightarrow 0$, we have $E_1 \rightarrow E_2 \rightarrow E''$ and $\left.\frac{\partial^2 y}{\partial t^2}\right|_{E_1} \rightarrow \left.\frac{\partial^2 y}{\partial t^2}\right|_{E_2} \rightarrow 0$. Consequently, the limit of the 0_2 solution when $q = 0$ is another static solution. This static solution has its own unique form of instability. When y is perturbed from its static value of $y = E''$, the solution ceases to exist. Consequently, we shall term this local behavior Einstein behavior of the second type and denote it as E_2. It could hardly be considered even as a rough approximation for the 0_2 solution. As shown in Table I, there is no solution at any point when $q \rightarrow 0$ from the other side. As before, these special

Table V

Λ	Conditions	$g(r)$	$h(r)$	Type	Range
+	$-g^3=\frac{9}{4}\Lambda h^2$	−	+	A_1	$0\leqslant y\leqslant E''$
				A_2	$E''\leqslant y\leqslant+\infty$
				E_1	$y=E''$
			−	M_2^x	$E\leqslant y\leqslant+\infty$
−	$g^3=-\frac{9}{4}\Lambda h^2$	+	−	E_2	$y=E''$
			+	O_1^x	$0\leqslant y\leqslant E$

conditions can also be satisfied with $g > 0$ and $h > 0$, and we obtain a non-elliptic special solution for the O_1 Solution G.

These six special solutions are summarized in Table V. The asterisks are used with M_1^* and O_1^* to indicate that these are special non-elliptic solutions. There are two static solutions, which will be considered in the following chapter. Here we shall proceed to derive the four non-static special solutions.

9. $\Lambda < 0$, $g > 0$, $h > 0$, O_1^*

Since our special condition is that $\alpha^3 = -\Lambda/48$, $\alpha(r)$ is a positive real constant which is entirely determined by the existing cosmological constant.

Assuming that this constant has been determined, the two periods are $\omega_1^x = \frac{\pi}{2\sqrt{3\alpha}} = \omega_1$ and $\omega_2^x = i\infty = \omega_2$ according to (3.35) and (3.36). The system of the period-parallelograms has degenerated into a vertical system of strips. There is only one path within the finite part of the ξ-plane which we can consider and that is $\xi = x$. Using the exact statement (3.39) we find the exact solution for this special case to be

$$y = \frac{3h}{2g}\left[\frac{3}{\cos 2\theta + 2} - 1\right] \tag{3.40}$$

$$t + f = \frac{3h}{g^{3/2}}\left[\sqrt{3}\,\tan^{-1}\left(\frac{1}{\sqrt{3}}\tan\theta\right) - \theta\right] \tag{3.41}$$

where $\theta = x\sqrt{3\alpha}$

This exact solution has been verified by substituting it into the original differential equation (2.1) under the special condition here assumed. This bit of mathematical formalism will not be presented here. This parametric solution may be seen by inspection to be cyclic with minimum values of $y = 0$ when $\theta = (n\pi)$ where n is any positive or negative integer. Similarly, the solution has maximum values of $y = 3h/g$ when $\theta = (n + \frac{1}{2})\pi$. The period of time in which one complete cycle would be executed would be $T = 3(\sqrt{3} - 1)\,h\pi/g^{3/2}$. This is O_1 behavior, as predicted.

10. $\Lambda > 0$, $g < 0$ $h > 0$, A_1

Here, $\alpha(r)$ is determined by the special condition as a real negative constant whose exact value is known, once the cosmological constant is found. Equations (3.35) and (3.36) then give for the two periods $\omega_1^x = -i\frac{\pi}{2\sqrt{-3\alpha}} = -\omega_2$ and $\omega_2^x = \infty = \omega_1$. Thus the period-parallelograms in the ξ-plane have here degenerated into a system of horizontal infinite strips. There are, accordingly, two systems of allowed lines for ξ. Let us first consider the real axis and all other congruent lines. Taking $\xi = x$ in (3.39) we find the exact solution

$$y = \frac{3h}{2(-g)}\left[1 - \frac{3}{\cosh 2\theta + 2}\right] \tag{3.42}$$

$$t + f = \frac{3h}{(-g)^{3/2}}\left[\frac{\sqrt{3}}{2}\log\left(\frac{\sqrt{3} - \tanh\theta}{\sqrt{3} + \tanh\theta}\right) + \theta\right] \tag{3.43}$$

where $\theta = x\sqrt{-3\alpha}$

This parametric solution gives $y = 0$ and $t + f = 0$ when $\theta = 0$ with y approaching $3h/2(-g)$ asymptotically and $t + f \to \pm\infty$ when $\theta \to \pm\infty$ This is the typical A_1 solution which was predicted in the earlier discussion.

11. $\Lambda > 0$, $g < 0$, $h < 0$, M^*_2

The second system of lines in the ξ-plane along with ξ can vary with physical significance is $\xi = \omega_2 + x$

and congruent lines. Taking this value of the parameter, we find

$$y = \frac{3(-h)}{2(-g)}\left[\frac{3}{2-\cosh 2\theta} - 1\right] \tag{3.44}$$

$$t+f = \frac{3(-h)}{(-g)^{3/2}}\left[\frac{\sqrt{3}}{2}\log\left(\frac{\coth\theta+\sqrt{3}}{\coth\theta-\sqrt{3}}\right) - \theta\right] \tag{3.45}$$

where $\theta = x\sqrt{-3\alpha}$ and $-0.659 \leqslant \theta \leqslant +0.659$.

As before, we have taken advantage of the $F(\xi_o)$ term in (2.8) to adjust the $t + f$ solution so that $t + f = 0$ when $\theta = 0$. The formal similarity of (3.43) and (3.45) with (2.8) might be noticed. When $\theta = 0$, we have the minimum value of y which is $3h/g$. As $|\theta|$ increases, we will have infinite values for both (3.44) and (3.45). In (3.44) we have $y = +\infty$ whenever $\cosh 2\theta = 2$. In (3.45) we have $t + f = +\infty$ whenever $\coth\theta = \sqrt{3}$ and $t + f = -\infty$ whenever $\coth\theta = -\sqrt{3}$. The same value of θ satisfies both of these conditions and is approximately $|\theta| = 0.659$. This is, of course, either χ or one of its congruent points. Thus these equations give a non-elliptic solution to the M_2 case under the stated special conditions.

12. $\Lambda > 0$, $g < 0$, $h > 0$, A_2

Continuing the solution for the remaining permitted variation of ξ, we find

$$y = \frac{3h}{2(-g)}\left[1 + \frac{3}{\cosh 2\theta - 2}\right] \tag{3.46}$$

$$t + f = \frac{3h}{(-g)^{3/2}}\left[\theta - \frac{\sqrt{3}}{2}\log\left(\frac{\sqrt{3} + \coth\theta}{\sqrt{3} - \coth\theta}\right)\right] \tag{3.47}$$

where $\theta = x\sqrt{-3\alpha}$ and $|\theta| > 0.659$.

These equations are merely (3.44) and (3.45) rearranged. An attempt has been made throughout this study to present the coefficients of bracketed terms as real positive quantities. Here, since $h(r)$ has changed signs between section 11 and 12, a sign has been altered in both the coefficient and the bracketed terms in stating (3.46) and (3.47). There are two domains of θ to be considered. When θ = ca. +0.659, $y = +\infty$ and $t + f = -\infty$ and while $\theta \to +\infty$, y approaches $3h/2(-g)$ asymptotically and $t + f \to +\infty$. Thus for this allowed range of θ, the local behavior of y with time is one which begins with infinite expansion and contracts asymptotically to the Einstein value. For the other range of θ, we have when θ = ca. -0.659 that $y = +\infty$ and $t + f = +\infty$, and when $\theta \to -\infty$ that y approaches $3h/2(-g)$ asymptotically with $t + f \to -\infty$. This local behavior of y with time is then the inverse of the preceding since here y begins with the Einstein value and expands from it asymptotically until it is infinite. This is, in both cases, the postulated A_2 behavior.

This last group of solutions is mathematically interesting since they are fully equivalent to the known Einstein and asymptotic solutions of homogeneous cosmology. However, they do not appeal too strongly to the physical intuition since they exist only under highly specialized conditions. If any assumptions are made as to the exact value of the cosmological constant, a more plausible assumption would seem to be that it is zero. This assumption, along with the special solutions that will arise with zero values for $g(r)$ and $h(r)$, will be considered in the chapter that immediately follows.

Chapter IV

Special Solutions and Conclusions

In previous chapters we have found the general solution to the partial differential equation (1.34) and have presented the approximations to the general solution under certain conditions. In this chapter we shall find all of the special solutions to the partial differential equation (1.34) when the coefficients are restricted to finite values. We shall summarize the various possible types of local behavior with a given value of the cosmological constant by presenting this material graphically. Finally, we shall derive a general solution for the proper local density within these models. We shall then see what further restrictions are placed upon our solutions by the physical requirement that the density of matter can never be negative. The material of this chapter will conclude the general theory of spherically-symmetric, non-homogeneous, relativistic models having zero pressure.

The general solution was derived in Chapter II when the three roots of the cubic in (2.2) were all distinct. A very obvious group of special solutions is found when two or more of these roots are repeated. The cubic will have two repeated roots whenever

$$\Lambda \left[\Lambda + 48 \alpha^3 \right] = 0 \tag{4.1}$$

This condition may be satisfied in two ways. If $\Lambda = 0$, condition (4.1) is satisfied and we have a repeated root of α and a single root of -2α. The condition (4.1) may also be satisfied with $\Lambda = -48\,\alpha^3$ and we will have a repeated root of $-\alpha$ and a single root of 2α.

Of these two alternates, the condition $\Lambda = 0$ is physically the most probable value. Einstein (Ref. 9, pp. 109-132) in a recent book has stated that he originally introduced the cosmological constant into his Field Equations only to make possible a static solution containing matter. While this term is logically permissible, he has observed that it makes for a considerable complication of the theory. To quote him exactly: "If Hubble's expansion had been discovered at the time of the creation of the general theory of relativity, the cosmological member would never have been introduced. It seems now so much less justified to introduce such a member into the field equations, since its introduction loses its sole original justification,--that of leading to a natural solution of the cosmologic problem."

Einstein and his pupils (e.g., Ref. 10) in recent papers have been using the Field Equations without the cosmological constant. Bergmann (Ref. 11, p. 179), an associate of Einstein, in a recent book states the Field

Equations without the cosmological constant, and no where in his text does he even suggest the possibility of such a term. Certainly, in deciding for or against a non-zero cosmological constant, the opinion of the founder of the relativity theory and of his school must be given due weight.

One of the most disturbing consequences of assuming a non-zero cosmological constant is that empty space is not flat, and that conversely, flat space is not empty. This has been noted by Gregory (Ref. 12, p. 180), among others. Tolman (Ref. 1, pp. 402-3) has ably summarized the matter. Quoting Tolman: " * * * $\Lambda = 0$ certainly seems the most reasonable assumption to make at the present time. In the first place the original argument, as discussed in TP139, for Einstein's addition of the logically permissible but otherwise surprising cosmological term to his original field equations in order to obtain a universe with a finite density of matter, now no longer exists in view of the wider possibilities presented by non-static models. In the second place, we have at the present time no accepted theory for any value at all for the cosmological constant, although interesting considerations concerning this matter have been presented by Eddington. And in the third place, from the observational point of view we can at least say that the value of Λ must be small in order not to upset the application of relativistic theory

to the orbits of the planets. Hence in what follows we shall lay special stress on the behaviour of models with the cosmological term omitted."

Taking $\Lambda = 0$ in equation (3.6) we have $q = 0$. Using $q = 0$ in the approximations of the first part of Chapter III, we obtain exact solutions for this special case. Only three types of solutions survive. We shall list these three solutions along with the conditions under which they exist.

1. $g(r) > 0 \quad h(r) > 0 \quad \Lambda = 0 \quad M_1$

$$y = \frac{h}{2g}\left[\cosh\theta - 1\right] \tag{4.2}$$

$$t + f = \frac{h}{2g^{3/2}}\left[\sinh\theta - \theta\right] \tag{4.3}$$

2. $g(r) < 0 \quad h(r) > 0 \quad \Lambda = 0 \quad O_1$

$$y = \frac{h}{2(-g)}\left[1 - \cos\theta\right] \tag{4.4}$$

$$t + f = \frac{h}{2(-g)^{3/2}}\left[\theta - \sin\theta\right] \tag{4.5}$$

3. $g(r) > 0 \quad h(r) < 0 \quad \Lambda = 0 \quad M_2$

$$y = \frac{(-h)}{2g}\left[\cosh\theta + 1\right] \tag{4.6}$$

$$t + f = \frac{(-h)}{2g^{3/2}}\left[\sinh\theta + \theta\right] \tag{4.7}$$

There is no solution in the fourth case in which $g(r) < 0$, $h(r) < 0$, and $\Lambda = 0$ since $\frac{\partial y}{\partial t}$ would be imaginary for all allowable values of $y(r,t)$.

It should be noted that the three preceding special solutions may be derived directly from the general solution (2.3) and (2.8), from the known degenerations of the Weierstrassian elliptic functions (vide, Ref. 6, pp. 105-6). In fact, the particular form of (2.8) was so chosen that this derivation might be performed without logarithmic difficulties. This derivation from the degenerations of the elliptic functions has been made, but the work will not be presented here. The derivation is long and tedious, but is no more rigorous than the obvious derivation from the material of Chapter III. Furthermore, the derivation by the degenerations of the elliptic functions yields no additional information about the special solutions. In addition, these three special solutions, as well as all other special solutions to be presented in this chapter, have been checked by direct substitution into the original partial differential equation. The procedure followed is obvious and will not be further commented upon.

The other alternative for a repeated root in the cubic of (2.2) is $\Lambda = -48\alpha^3$. The six special solutions resulting from this condition have been discussed in the last part of Chapter III. The actual solutions for the

four non-static cases will be found there and need not be repeated in this chapter. The six special solutions and the conditions under which they exist are:

1. $g(r) > 0$ $h(r) > 0$ $\Lambda = -48\alpha^3$ O_1^*

Equations (3.40) and (3.41)

2. $g(r) < 0$ $h(r) < 0$ $\Lambda = -48\alpha^3$ M_2^*

Equations (3.44) and (3.45)

3. $g(r) < 0$ $h(r) > 0$ $\Lambda = -48\alpha^3$ A_1

Equations (3.42) and (3.43)

4. $g(r) < 0$ $h(r) > 0$ $\Lambda = -48\alpha^3$ A_2

Equations (3.46) and (3.47)

5. $g(r) < 0$ $h(r) > 0$ $\Lambda = -48\alpha^3$ E_1

This case was discussed in Chapter III, but the solution was not given. It may readily be seen that the solution in this case is

$$y = -\frac{3h}{2g} \tag{4.8}$$

where t can take any real value. The condition $\Lambda = -48\alpha^3$ defines a functional relationship between $g(r)$ and $h(r)$. Substituting this functional relationship and solution (4.8) into the partial differential equation (1.34), we see that it is satisfied identically and hence that $\frac{\partial y}{\partial t} = 0$. Equation (1.34) can be differentiated to give an expression for $\frac{\partial^2 y}{\partial t^2}$ in terms of $y(r,t)$, $g(r)$, $h(r)$, and Λ.

Substituting the functional relationship and (4.8) into this derived equation, we see that it also is satisfied identically and that $\frac{\partial^2 y}{\partial t^2} = 0$. Hence from the reasoning of Chapter III it follows that all higher partial time derivatives are also identically zero. Consequently, equation (4.8) is a static, albeit unstable, solution to the problem.

If we take our functions as $g(r) = - r^2/R^2$ and $h(r) = 2r^3/3R^2$, where R is a real positive constant, then (4.8) reduces to $y(r) = r$. Substituting this into equation (1.23) gives the corresponding line-element as

$$ds^2 = -\left[\frac{dr^2}{1 - \frac{r^2}{R^2}} + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2\right] + dt^2 \qquad (4.9)$$

Substituting the same values into equation (1.36) gives the density as

$$8\pi\rho = \frac{2}{R^2} \qquad (4.10)$$

While a substitution into the condition $\Lambda = - 48\alpha^3$ gives

$$\Lambda = \frac{1}{R^2} \qquad (4.11)$$

Equations (4.9), (4.10), and (4.11) are, respectively, the line-element, the density, and the curvature of a static Einstein model with zero pressure. Furthermore, any other E_1 solution can be converted into this form by a simple

transformation of the r-coordinate. It is for this reason that the E_1 solutions have been designated as "Einstein behavior of the first type".

6. $g(r) > 0 \quad h(r) < 0 \quad \Lambda = -48\alpha^3 \quad E_2$

These conditions are those of the previous case with all signs changed. The solution to these conditions is still equation (4.8). However, if the attempt is made to extend the E_2 solution over a finite range of r, it will be found from equation (1.36) that the resulting density is negative. Hence the E_2 solution does not exist over a finite volume of space, but only for isolated values of r. It is a transition case between two non-static solutions. For these reasons the E_2 solutions have been designated "Einstein behavior of the second type".

We know from the homogeneous theory there are only three possible static models, of which only one, the Einstein model, contains matter. The two remaining static models, namely the de Sitter and the Special Relativity solutions, should be deductible as special cases of the non-homogeneous theory. Although it is somewhat auxiliary to the present line of development, they might best be considered at this time.

1. $g(r) = 0 \quad h(r) = 0 \quad \Lambda = 0$ Special Relativity

The partial differential equation (1.34) is readily integrated under these special conditions as

$$y(r) = f(r). \tag{4.12}$$

If we take $y(r) = f(r) = r$, then the line-element becomes

$$ds^2 = -\left[dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right] + dt^2 \tag{4.13}$$

which is the line-element of Special Relativity. Since $h(r)$ = constant, this model does not contain matter.

2. $g(r) = 0$ $h(r) = 0$ $\Lambda > 0$ de Sitter

If we take $\frac{1}{3}\Lambda = k^2 = 1/R^2$, the partial differential equation (1.34) can be integrated as

$$y(r) = e^{\kappa(t+f)} \tag{4.14}$$

If we define the arbitrary function as $f(r) = R \log r$, then we have $y(r) = r\, e^{kt}$ and the line-element becomes

$$ds^2 = -e^{2kt}\left[dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right] + dt^2 \tag{4.15}$$

This is the Lemaitre (Ref. 13, p. 188) - Robertson (Ref. 14, p. 835) form of the de Sitter line-element. It can be transformed into the more conventional form of the de Sitter line-element by the transformation equations given by Tolman (Ref. 1, eq. 142.8, p. 347). Here again, since $h(r)$ = constant, the model contains no matter.

To return to the special solutions of the partial differential equation (1.34) when the cubic in (2.2) has repeated roots, we have one more case to consider. This is

the case when all three roots are repeated. The cubic of (2.2) will have a triply-repeated root of zero when $g_2 = 0$ and $g_3 = 0$. The cubic of (2.3) cannot have a triply-repeated non-zero root because of the absence of the y^2 term. The conditions can be satisfied only by $\alpha = 0$ and $\Lambda = 0$. The condition $\alpha = 0$ can be satisfied either by taking $g(r) = 0$ or by taking $h(r) = \infty$. However, we wish to restrict our solutions to those arising from finite coefficients in equation (1.34); hence we must reject $h(r) = \infty$ and take instead $g(r) = 0$. Under these conditions the differential equation may be integrated explicitly as

$$y(r,t) = \left(\frac{9}{4} h\right)^{1/3} (t+f)^{2/3} \qquad (4.16)$$

The conditions under which this solution exists are:

$g(r) = 0$	$h(r) > 0$	$\Lambda = 0$	M_1

The factor of $9/4$ which is the coefficient of $h(r)$ in this solution is actually redundant since it could be absorbed into the arbitrary function. However, the general solution (2.2) and (2.8) reduces exactly to (4.16) with the degenerations of the Weierstrassian elliptic functions when all three roots are equal. The special solution (4.16) is an extension to all values of r and t of the known behavior of the general O_1 and M_1 solutions in the neighborhood of $y = 0$.

The foregoing discussion does not include all of the special solutions to the partial differential equation (1.34). When one or more of the coefficients of this equation, namely $g(r)$, $h(r)$, and Λ, are zero, the degree of the polynomial may be reduced by transformations or otherwise and the special solution found in terms of non-elliptic functions. All of the possible special solutions when the three coefficients are limited to finite values are listed in Table VI. Since most of these solutions have already been discussed, they are tabulated in the table by equation numbers only. The remaining nine solutions which have not been discussed are given in Table VI along with the conditions under which they exist. These solutions are readily derived by elementary mathematics in each case and do not require discussion here. Two limiting cases might be noted. They are $y = \infty$, which is denoted by I for "infinite model"; and $y = 0$, which is denoted by Z for "zero model". These are themselves limiting cases of other special solutions.

Table VI and Table I give all of the solutions to the partial differential equation (1.34) when the three coefficients have finite values. This tabular material might be made more comprehensible by presenting the conditions under which the various solutions exist in

Table VI

Λ	$g(r)$	$h(r)$	Solution	Symbol
0	+	+	eqs. (4.2) and (4.3)	M_1
	−	+	eqs. (4.4) and (4.5)	O_1
	+	−	eqs. (4.6) and (4.7)	M_2
	−	−	no solution	n.s.
$-48\alpha^3$	+	+	eqs. (3.40) and (3.41)	O_1^x
	+	−	eq. (4.8)	E_2
	−	−	eqs. (3.44) and (3.46)	M_2^x
	−	+	eqs. (3.42) and (3.43)	A_1
			eqs. (3.46) and (3.47)	A_2
			eq. (4.8)	E_1
+	0	+	$y=\left(\frac{3h}{\Lambda}\right)^{1/3}\sinh^{2/3}\frac{\sqrt{3\Lambda}}{2}(t+f)$	M_1
−		+	$y=\left(\frac{3h}{-\Lambda}\right)^{1/3}\sin^{2/3}\frac{\sqrt{-3\Lambda}}{2}(t+f)$	O_1
+		−	$y=\left(\frac{-3h}{\Lambda}\right)^{1/3}\cosh^{2/3}\frac{\sqrt{3\Lambda}}{2}(t+f)$	M_2
−		−	no solution	n.s.
+	+	0	$y=\sqrt{\frac{3g}{\Lambda}}\left\lvert\sinh\sqrt{\frac{\Lambda}{3}}(t+f)\right\rvert$	M_1
−	+		$y=\sqrt{\frac{3g}{-\Lambda}}\left\lvert\sin\sqrt{\frac{-\Lambda}{3}}(t+f)\right\rvert$	O_1
+	−		$y=\sqrt{\frac{-3g}{\Lambda}}\cosh\sqrt{\frac{\Lambda}{3}}(t+f)$	M_2
−	−		no solution	n.s.
0	+	0	$y=\sqrt{g}\,\lvert t+f\rvert$	M_1
	0	+	eq. (4.16)	M_1
	−	0	no solution	n.s.
	0	−	no solution	n.s.
	0	0	eq. (4.12)	S.R.
	0	+	$y=\infty$	I
+	0	0	eq. (4.14)	d.S.
±	0	0	$y=0$	Z

graphical form. This is a logical extension of Robertson's (Ref. 4) similar treatment for the homogeneous models.

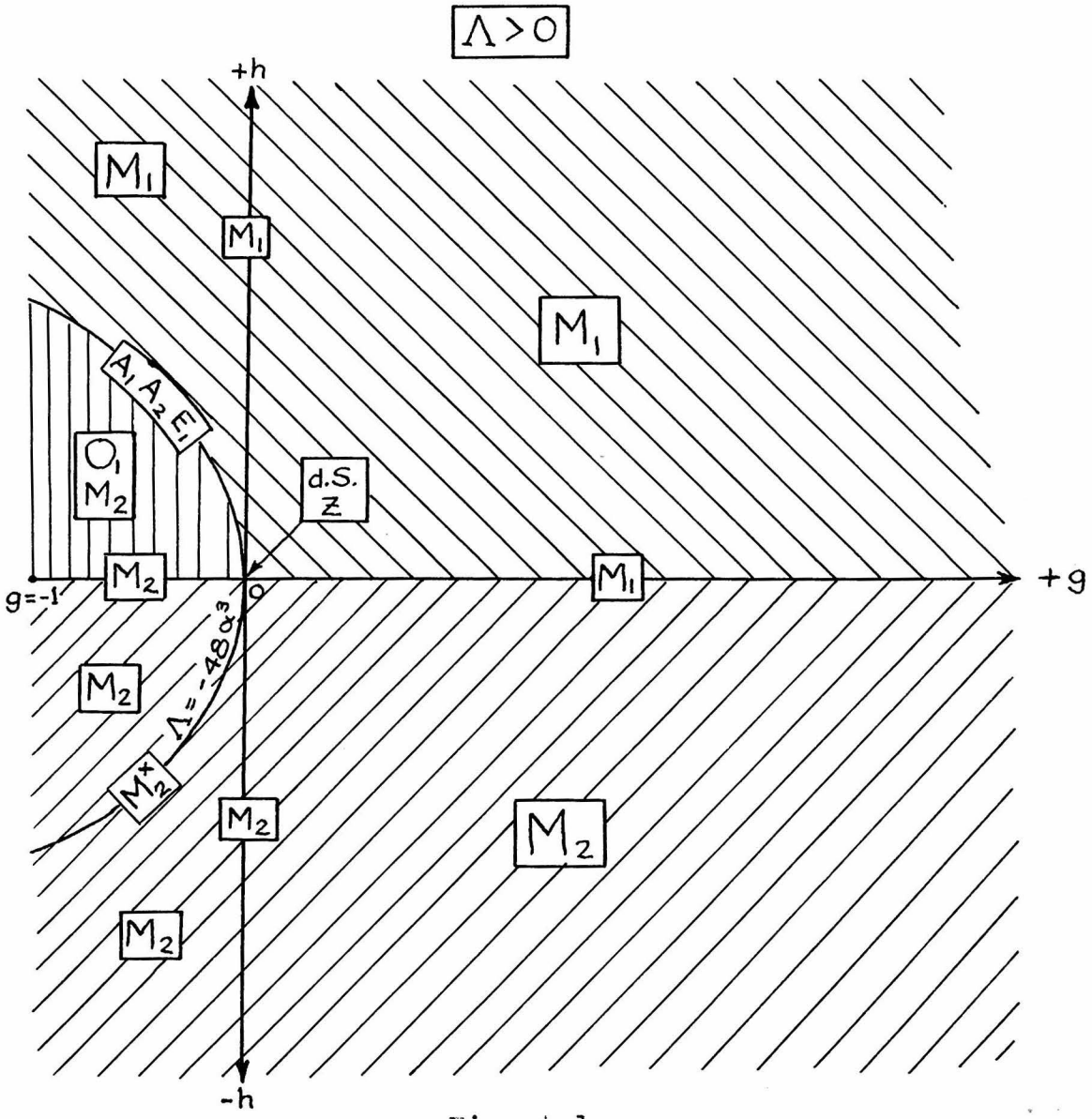

Fig. 4.1

Robertson defined a certain critical quantity which he denoted as Q and plotted this critical quantity against R, a measure of the instantaneous proper distances within

the model. Robertson made two separate plots over half-planes; one for the open models, and one for the closed

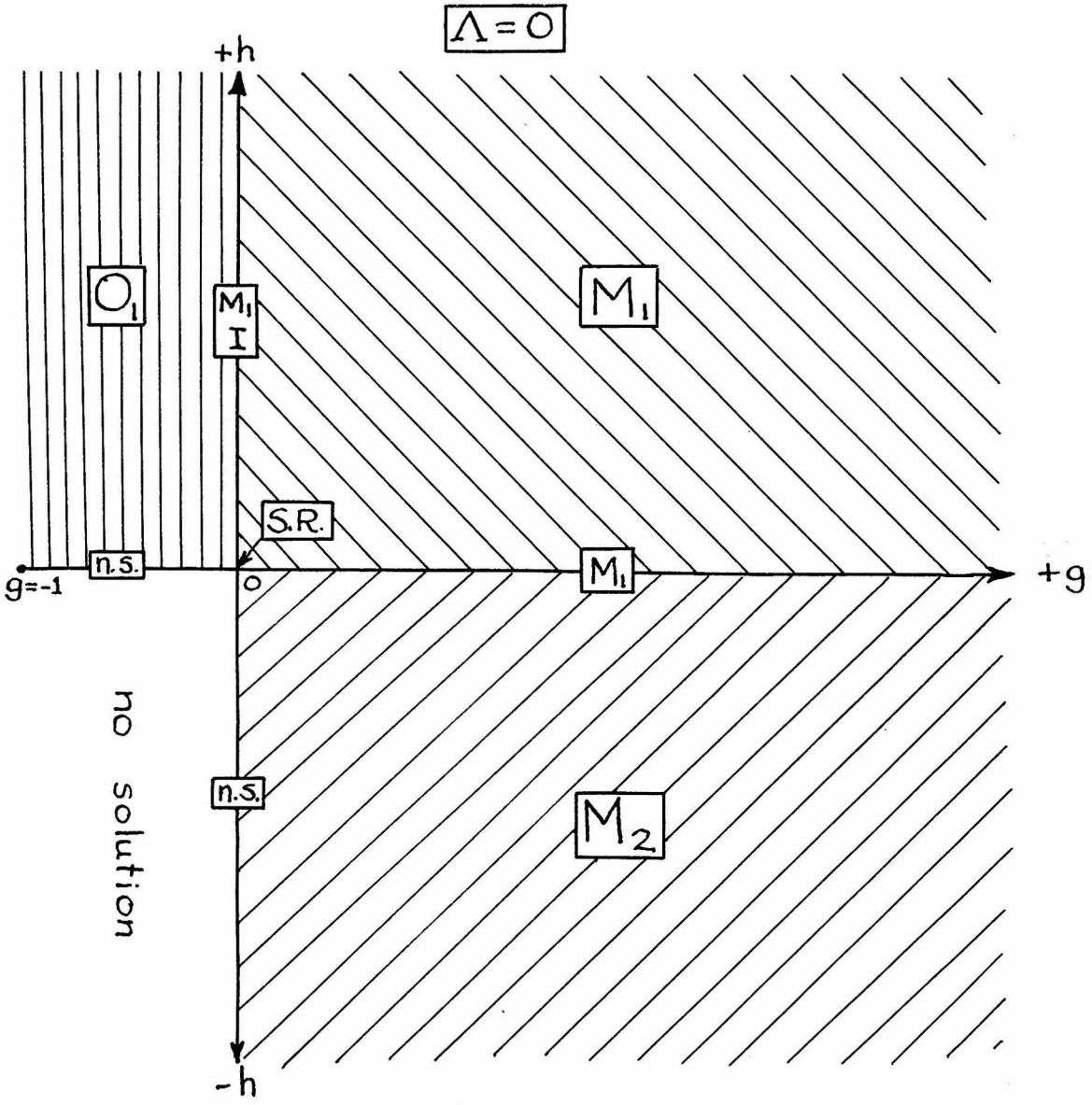

Fig. 4.2

models. By the obvious artifice of plotting Q against R^2, these two separate plots can be combined into a single curve over the entire Q - R^2 plane. This plot of Q against

R^2 can be interpreted as delineating areas within the $q - R^2$ plane in which the various possible homogeneous

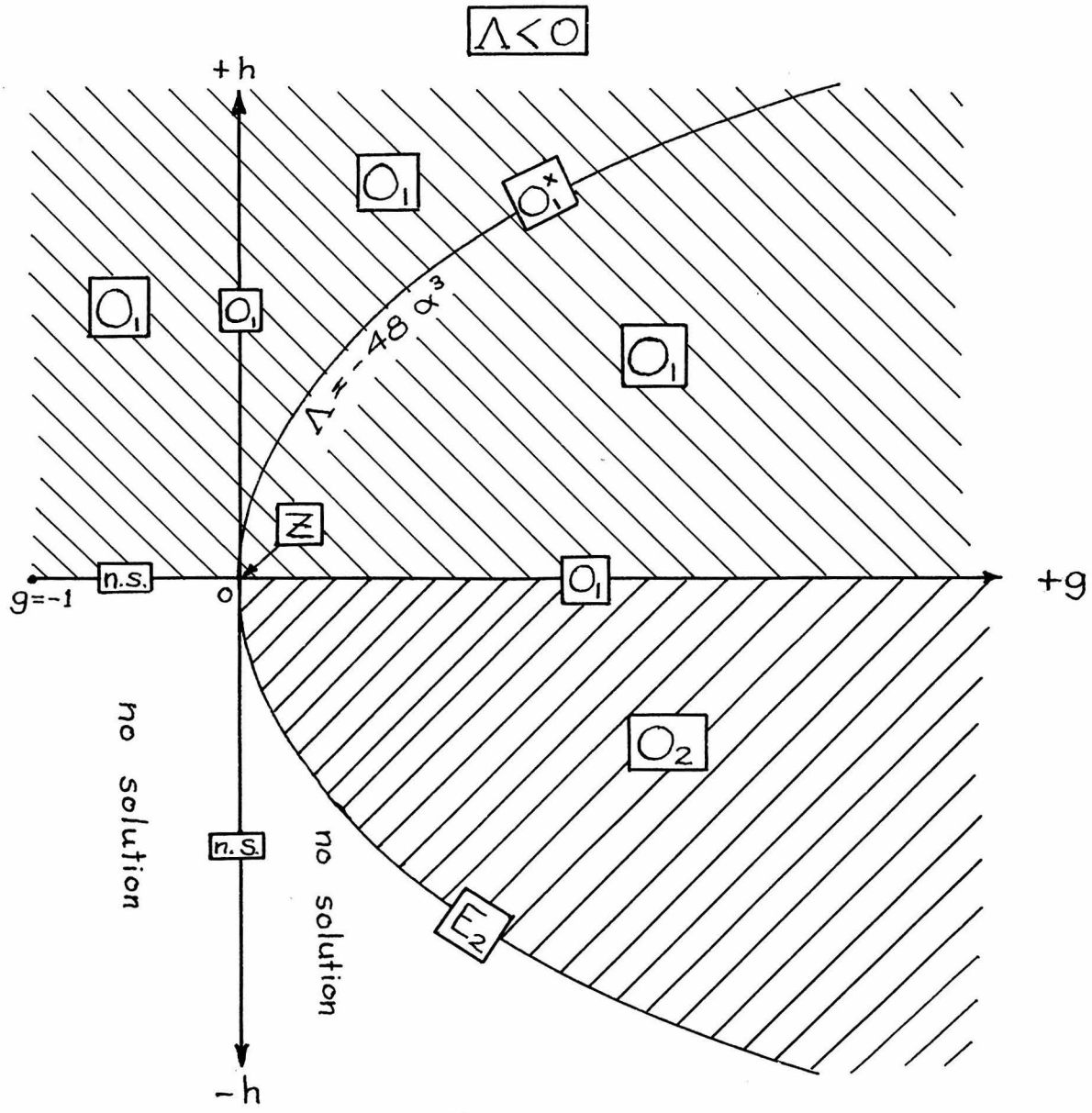

Fig. 4.3

solutions would exist. The solutions which will exist for a particular value of Λ are then found as a one-dimensional domain, formed by the insection of a straight

line (Q = constant) with these areas.

Since the non-homogeneous models are a higher order of generalization than the homogeneous models, a graphical presentation of the solution conditions must be made in three or more dimensions. The method that we shall follow is to plot the data of Tables I and VI in three-dimensional $g - h - \Lambda$ space. The general solutions of Table I exist within volumes delineated by the planes $g(r) = 0$, $h(r) = 0$, $\Lambda = 0$, and by the surface $\Lambda = -48\alpha^3$. The special solutions of Table VI exist within areas upon these planes or surface. Certain limiting forms of the special solutions in Table VI exist only along lines at isolated points within the $g - h - \Lambda$ space. The solutions which will exist for a particular value of Λ are then found as a two-dimensional domain by the intersection of a plane (Λ = constant) with the various volumes, areas, lines or points.

Since it is manifestly difficult to give a clear representation of a three-dimensional plot upon two-dimensional paper, the various relations are indicated in the accompanying figures as representative intercepts of three different Λ = constant planes with the $g - h - \Lambda$ space. Fig. 4.1 shows a typical intersection when $\Lambda > 0$. The general solutions of Table I which will apply in this case are shown as areas in Fig. 4.1 which have been cross-

hatched to distinguish the various types of local behavior. Each of these areas represents a single type of local behavior, except for one area in the second quadrant. Within this exceptional area the local behavior may be either O_1 or M_2, depending on the particular range of $y(r,t)$ which is involved. This is in accordance with Table I and the two allowed ranges for $y(r,t)$ are non-overlapping. The special solutions of Table VI are shown in Fig. 4.1 as existing along lines or at a point at the origin. The A_1, A_2, and E_1 special solutions exist along the upper-half of the curve $\Lambda = -48\alpha^3$, while the non-elliptic M_2^* solution exists along the lower-half of this curve. This curve approaches the line $g = 0$ as $\Lambda \to 0$, and approaches the line $0 \geqslant g \geqslant -1$ as $\Lambda \to \infty$. Only a half-plane is represented in Fig. 4.1 since we must have $g(r) \geqslant -1$. The approximations of the first part of Chapter III apply whenever $|\Lambda| < |\alpha^3|$. Thus, for a finite Λ, there is always an area in Fig. 4.1 in which these approximations would apply. This area approaches the line $h(r) = 0$ when $\Lambda \to +\infty$. On the other hand, this area fills more and more of the g - h plane as $\Lambda \to 0$ becoming equal to the entire plane when $\Lambda = 0$. Any particular non-homogeneous model could be diagrammed as a curve on the g - h plane in Fig. 4.1 which includes all

of the values assumed by $g(r)$ and $h(r)$. This curve could include any or all of the possible local types of behavior shown in Fig. 4.1.

The transition case $\Lambda = 0$ is shown in Fig. 4.2. Here all of the solutions are non-elliptic. The curve $\Lambda = -48\alpha^3$ has degenerated into the line $g(r) = 0$. A new feature appears in the third quadrant. No physical solutions are possible with the $g(r)$ and $h(r)$ values in the third quadrant. The point at the origin represents the limiting special case of the Special Relativity solution. The remaining case of $\Lambda < 0$ is shown in Fig. 4.3. The area of no physical solution has widened to include part of the fourth quadrant. As $\Lambda \to -\infty$, this area continues to increase until in the limit there is no physical solution within both the third and the fourth quadrants. The curve $\Lambda = -48\alpha^3$ has shifted to the right-hand side of the line $g(r) = 0$. As $\Lambda \to -\infty$, this curve approaches more and more closely to the line $0 \leqslant g \leqslant +\infty$. The special solution E_2 exists along the lower half of this curve, while the non-elliptic solution O_1^* exists along the upper half. As before, the approximations of the first part of Chapter III apply to a symmetrical area lying about the line $h(r) = 0$. When $\Lambda = 0$, this area of approximation covers the entire g - h plane, but as $\Lambda \to -\infty$, this area approaches more

and more closely to the line $h(r) = 0$.

Thus far in this study we have restricted our solutions to those which will yield positive values for $y(r,t)$ and real values for t. Physically, these are necessary but not sufficient conditions. A physically crucial condition is that the local proper density at all times and all places within the model must either be zero or a real positive value. A negative density would have no physical meaning. We shall devote the remainder of this chapter to deriving a general expression for the proper local density within the models and to a consideration of further restrictions which might be placed upon the solutions summarized in Tables I and VI by the requirement that the density never be negative.

The proper local density may be found by substituting the general solution (2.3) into the expression for the density (1.36). This yields

$$8\pi\rho = \frac{12\left[\beta(\xi) - \alpha\right]^3 \frac{dh}{dr}}{\frac{dh}{dr} - 3h\frac{\beta'(\xi)}{\beta(\xi)-\alpha}\frac{\partial\xi}{\partial r} - 3h\frac{\frac{\partial}{\partial\alpha}\{\beta(\xi)-\alpha\}}{\beta(\xi)-\alpha}\frac{d\alpha}{dr}} \tag{4.17}$$

We wish to express (4.17) in a form which depends only on the local values of $g(r)$, $h(r)$, ξ, and the constant value of Λ. The derivative $\frac{d\alpha}{dr}$ is acceptable, since $\alpha(r) = g/3(2h)^{2/3}$, and so this derivative can be expressed

in terms of $g(r)$ and $h(r)$ only. The partial derivative $\frac{\partial \xi}{\partial r}$ is not so amiable. It is not immediately obvious how ξ varies with r and t. A relationship for $\frac{\partial \xi}{\partial r}$ can be found by differentiating equation (2.5), the second member of the general solution pair, with respect to r. This gives

$$\frac{df}{dr} = \frac{\partial}{\partial \alpha} \int_{\xi_o}^{\xi} \frac{du}{\wp(u) - \alpha} \frac{d\alpha}{dr} + \frac{1}{\wp(\xi) - \alpha} \frac{\partial \xi}{\partial r} \tag{4.18}$$

where ξ_o is the value of the parameter when $t + f = 0$ and ξ is the present value of the parameter for the particular values of r and t, which are being considered. The derivative $\frac{\partial \xi}{\partial r}$ can now be eliminated by combining equations (4.17) and (4.18).

To carry out the indicated procedure, an expression must be obtained for the derivative $\frac{\partial}{\partial \alpha} \wp(\xi ; \alpha)$. The partial derivative of the Weierstrassian Pe-function with respect to its invariants cannot be found in any of the standard mathematical works, but can be derived with some labor. The steps in this derivation will only be sketched out herein. We begin with the inverse elliptic integral, equation (2.10), which corresponds to the Weierstrassian Pe-function. First, we must differentiate this integral with respect to $\alpha(r)$. The resulting

expression contains a new elliptic integral which may be evaluated in the usual way by expressing the integrand as a sum of Zeta-functions and thereby carrying out the indicated integration. This result will have coefficients involving $\alpha(r)$ and the three roots, e_1, e_2, and e_3. This is not the most useful form. However, this expression may be further reduced, with some patience, by the known relations between $\alpha(r)$, e_1, e_2, and e_3 and the invariants, g_2 and g_3, of the Weierstrassian Pe-function. The result of this work is

$$\frac{\partial}{\partial\alpha}\wp(\xi;\alpha) = -\frac{\alpha}{4\left[\alpha^3+\frac{1}{48}\Lambda\right]}\left[\wp'(\xi;\alpha)\left\{\zeta(\xi;\alpha)-\alpha\xi\right\}+2\left[\wp(\xi)-2\alpha\right]\left[\wp(\xi)+\alpha\right]\right] \qquad (4.19)$$

The validity of (4.19) can be checked in at least two ways. It is readily verified in the neighborhood of $\xi = 0$. The series expansion of the Weierstrassian Pe-function (vide Ref. 6, p. 99) is uniformly convergent and may be differentiated term by term. This gives a series expansion for $\frac{\partial}{\partial\alpha}\wp(\xi;\alpha)$. Equation (4.19) may be expanded into a similar series and the coefficients of the two series may be compared. This has been done for the first few terms and they were found to be identical. The second method is to derive an approximate expression for $\frac{\partial}{\partial\alpha}\wp(\xi;\alpha)$ when $|\Lambda| \ll |\alpha^3|$. This is readily done by a transformation of the invariants of the Weierstrassian Pe-function.

An approximation to (4.19) under the same conditions will be found to yield the same expression. Thus the derivation of (4.19) would seem to be without error.

Substituting (4.19) into (4.17), we have a complete expression for the last term in the denominator. Substituting (4.19) into (4.18), we have an expression which contains an integral which may again be integrated by the use of certain artifices. The partial derivative may be eliminated between the two resulting equations and we find the general solution for the proper local density to be

$$8\pi\rho = \frac{12\left[\rho(\xi)-\alpha\right]^3 \frac{dh}{dr}}{\frac{dh}{dr} - 3h\rho'(\xi)\,A + \frac{3h}{4\left[\alpha^3+\frac{1}{48}\Lambda\right]}\,B\,\frac{d\alpha}{dr}} \tag{4.20}$$

where $$A = \left[\frac{df}{dr} - \frac{\alpha}{4\left\{\alpha^3+\frac{1}{48}\Lambda\right\}}\left\{\frac{1}{2\alpha}\,\zeta(\xi_0) + \frac{1}{4\alpha}\,\frac{\rho'(\xi_0)}{\rho(\xi_0)-\alpha} + \frac{\zeta(\xi_0)-\alpha\xi_0}{\rho(\xi_0)-\alpha}\right\}\frac{d\alpha}{dr}\right]$$

and $$B = \left[\frac{\rho'(\xi)}{2}\left\{(\xi-\xi_0)\alpha - \zeta(\xi)\right\} - \left\{\rho(\xi)-2\alpha\right\}\left\{\rho(\xi)+\alpha\right\}\right]$$

There are several interesting cases of the general solution for the proper local density. In the special case that $\alpha(r)$ = constant and $f(r)$ = constant, we have

$$8\pi\rho = 12\left[\rho(\xi)-\alpha\right]^3 \tag{4.21}$$

Furthermore, under these conditions, it can readily be shown that ξ is a function of t alone since all of the

partial derivatives of ξ with respect to r can be shown to be zero. Hence, equation (4.21) is also a function of t only under these conditions and the model has been reduced to the non-static, homogeneous special case of the general solution. These homogeneous special cases will have physical validity only when the density, as given by equation (4.21), is greater or equal to zero. This requires that $[\rho(\xi) - \alpha] \geqslant 0$, which will be true for all solutions in which $h(r) > 0$. This excludes only one type of local behavior listed in Table I, namely the O_2 local behavior of solution I. Thus it is possible to have homogeneous special cases having O_1, M_1, and M_2 behavior or having any further specialization of these, such as A_1, A_2, E_1, S.R., d.S., I, or Z; but it is not possible to have a homogeneous model with O_2 or E_2 behavior as a special case of our general solutions (2.3) and (2.8).

One possible choice for $g(r)$ and $h(r)$ which will satisfy the condition $\alpha(r) =$ constant, is $g(r) = -r^2/R^2$ and $h(r) = 4k^3 r^3$, where k is a scale-constant which will be used to adjust the co-moving coordinate system so that it agrees with proper measure at a particular time, say when $t = t_o$. Using these values for $g(r)$ and $h(r)$, we find that $y = \frac{kr}{\rho(\xi) - \alpha}$. As before, if we take $f(r) =$ constant, then $[\rho(\xi) - \alpha]$ is a function of t alone.

To make the co-moving coordinate system agree with proper measure at $t = t_0$ we must have $y = r$ at the time; hence we must take $k = [\rho(\xi) - \alpha] = \left(\frac{2}{3}\pi\rho_0\right)^{1/3}$ where ρ_0 is the density at $t = t_0$. Using these values in (1.35), the line-element may be expressed as

$$ds^2 = -e^{g(t)}\left[\frac{dr^2}{1-\frac{r^2}{R^2}} + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\right] + dt^2 \qquad (4.22)$$

where $e^{g(t_0)} = 1$. This is one of the well-known forms of the homogeneous non-static line-element. We may now summarize the discussion by stating that when $\alpha(r) =$ constant, $f(r) =$ constant, and $h(r) > 0$, the general solutions (2.3) and (2.8) degenerate into homogeneous solutions having zero pressure.

The existence of a sub-class of homogeneous solutions suggests the further exploration of all possible static solutions. It was shown earlier in this chapter that the only static solution which contains matter and which applies over a finite volume of space is the E_1 solution. This solution exists, subject to the restriction that $48\alpha^3 = 4g^3/9h^2 = -\Lambda =$ - constant. If the coordinate system used is to be equal to the proper measure, then we must also have $h(r) = -3h/2g = r$. Solving these two conditions together, we find that we must have $h(r) = (2/3)\, r^3$ (constant). Since we can denote the constant

as we please, we shall take $h(r) = 2r^3/3R^2$, from which it immediately follows that $g(r) = -r^2/R^2$ and $8\pi\rho = 2/R^2$. Thus the only static solution of physical interest is homogeneous. It thereby follows that there are no static non-homogeneous models which contain matter and which apply over a finite volume of space.

The conditions $\alpha(r) = \text{constant}$ and $f(r) = \text{constant}$ were rather drastic. Let us try the less restrictive conditions that $\alpha(r) = \text{constant}$ but that $f(r) \neq \text{constant}$. Then equation (4.20) reduces to

$$8\pi\rho = \frac{12\left[\wp(\xi) - \alpha\right]^3}{1 - \frac{3h}{\frac{dh}{dr}}\wp'(\xi)\frac{df}{dr}} \tag{4.23}$$

which is a non-static, non-homogeneous statement for density. It is obvious that O_1, M_1, or other solutions in which the parameter includes $\xi = 0$ cannot exist under these conditions. The elliptic function $\wp'(\xi)$ has a triple pole at $\xi = 0$, and so we would have $\rho = (\text{constant})/\xi^3$ in the neighborhood of $\xi = 0$. This would require the existence of negative densities, so these solutions are not acceptable. It is also obvious that no solutions can exist subject to these conditions with $h(r) < 0$ since $\wp'(\omega_i) = 0$ and negative densities would once more appear within the model. Thus the only solutions which can exist subject to these conditions are the M_2 behavior of solution C of Table I,

or one of its special cases such as the A_2 solution. Furthermore, these solutions will exist only if $3h\rho'(\xi)\frac{df}{dr} < \frac{dh}{dr}$.

The most general solution (4.20) can be somewhat simplified by considering the solutions in groups according to the value of ξ_0. For O_1, M_1, A_1, etc. solutions in which we can take $\xi_0 = 0$, this expression becomes

$$8\pi\rho = \frac{12\left[\rho(\xi)-\alpha\right]^3\frac{dh}{dr}}{\frac{dh}{dr} - 3h\rho'(\xi)\frac{df}{dr} + \frac{3h}{4\{\alpha^3+\frac{1}{48}\Lambda\}}\left[\frac{\rho'(\xi)}{2}\{\alpha\xi - \zeta(\xi)\} - \{\rho(\xi)-2\alpha\}\{\rho(\xi)+\alpha\}\right]} \quad (4.24)$$

If in (4.24) we have $h\, df/dr \neq 0$, then in the neighborhood of $\xi = 0$ we again have $\rho = (\text{constant})/\xi^3$ and the appearance of negative densities. Therefore, a mathematical condition on these solutions is that $f(r) = \text{constant}$. With $f(r) = \text{constant}$ there is no difficulty since the last group of terms in the denominator will vary as ξ^2 in the neighborhood of $\xi = 0$, and so the density will vary as $\rho = (\text{constant})/\xi^6$ in this neighborhood and accordingly, will remain positive at all times and places.

The O_1 solution is cyclic. While it will behave properly in the neighborhood of the singularity at $\xi = 0$ with the above condition, this will not be so for the congruent singularities at $\xi = 2n\omega_1$. This follows since it is possible to prove that $\zeta(\omega_1) \neq \alpha\omega_1$ and hence the

density will again vary as $\rho = (\text{constant})/\xi^3$ in the neighborhoods of the singularities at $\xi = 2n\omega_1 \quad (n \neq 0)$ where $\wp'(\xi)$ has a triple pole. Thus the additional restriction of $\alpha(r) = \text{constant}$ must be imposed upon the O_1 solutions. This means that the only allowable O_1 solutions are the homogeneous models.

These restrictions upon the O_1 solutions are probably more mathematical than physical. The basic assumption of zero pressure is certainly violated in the neighborhoods of the singularities and so our equations are actually non-applicable. However, it would also seem impossible to have a physical condition in which the densities varied cyclically at all points, approaching infinity once each cycle, but where the period of the cycles is a non-constant function of r. This would require after two or three cycles from $\xi = 0$, that the density at some point, say r_o, would be approaching infinity, while the density at another point which is separated from r_o only by an infinitesimal distance would be small. This would entail a type of discontinuity which is not to be expected in macroscopic physics. One would expect that the effect of the neglected pressure term would be to slow the rate of contraction in the more rapidly contracting regions so that the true period of the cycle would be the same at all points of the model. Therefore the O_1 solutions derived in this

study should be regarded as only approximations to the true O_1 behavior where the approximation is good only over a part of one cycle.

The same remarks apply to the condition $f(r) =$ constant. This is a mathematical rather than a physical restriction. However, the M_1 and the A_1 solutions spend only a vanishingly small part of their history in the epoch in which pressure terms are important. Consequently, the M_1 and A_1 solutions derived herein should be very good approximations to the true behavior except in the immediate neighborhood of the singularity.

In the remaining solutions we may take $\xi_o = \omega_i$ where ω_i is one of the half-periods of the elliptic functions. Equation (4.20) then becomes

$$8\pi\rho = \frac{12\left[\wp(\xi)-\alpha\right]^3 \frac{dh}{dr}}{\frac{dh}{dr} - 3h\wp'(\xi)A + \frac{3h}{4\left[\alpha^3+\frac{1}{48}\Lambda\right]} B \frac{d\alpha}{dr}} \qquad (4.25)$$

where $$A = \left[\frac{df}{dr} - \frac{\alpha}{4\left\{\alpha^3+\frac{1}{48}\Lambda\right\}} \frac{\zeta(\omega_i) - \alpha\omega_i}{e_i - \alpha}\right]$$

and $$B = \left[\frac{\wp'(\xi)}{2}\left\{\alpha(\xi-\omega_i) - \zeta(\xi-\omega_i)\right\} + \left\{\wp(\xi)-e_j\right\}\left\{\wp(\xi)-e_k\right\} - \left\{\wp(\xi)-2\alpha\right\}\left\{\wp(\xi)+\alpha\right\}\right]$$

and where $e_i = \wp(\omega_i)$, while e_j and e_k are the remaining two roots. The last term in quantity A can be shown to be a pure imaginary when $\xi_o = \omega_2$. This term can be shown to have a complex value whenever $\xi_o = \omega_3$. There are no compensating terms elsewhere in (4.25). In

particular, the quantity B has only real values throughout the parameter range of ξ . This complex quantity A would produce a complex density in (4.25) which would have no physical meaning whatsoever. Therefore, we must conclude that solutions whose parameter ranges include either of the points ω_2 or ω_3 have no physical existence unless we also have $\alpha(r)$ = constant. This latter possibility has already been explored and has been found to yield a single M_2 solution (solution C of Table I). Nothing is gained by trying to take another point within the parameter range for ξ_o since the complex values for density will still be found.

This is an important conclusion, because it shows that the O_2 solution (solution I of Table I) and one M_2 solution (part of solution F in Table I) do not exist. Neither do their special forms. Thus the strange E_2 solution which was a special case of the O_2 has no actual existence. This conclusion puts the non-homogeneous theory into strict analogy with the homogeneous theory. The O_2 solution can exist in the homogeneous theory only through the presence of a particular type of pressure. Homogeneous models with zero pressure will not have O_2 solutions. Thus the possible types of local behavior within the non-homogeneous models which we have developed are identical to the possible types of behavior within homogeneous models having zero

pressure. No new types of local behavior have been produced by the introduction of a non-homogeneous distribution of matter within the model.

Returning to equation (4.25), if we take $\xi_o = \omega_1$ the last term in the quantity A can be shown to be real. Thus this M_2 solution is possible. The value of the quantity A is fixed for a given value of r. Hence the value of the collection of terms $3\,h\,\rho'(\xi)\,A$ is odd with respect to ω_1. The value of the quantity B varies with the parameter ξ, but is even in value with respect to ω_1. Therefore, the variation in density at a given point, say $r = r_o$, is not in general symmetric with respect to $t + f = 0$. The variation in a proper distance perpendicular to the r-coordinate at $r = r_o$, however, will be strictly symmetrical with respect to $t + f = 0$. This seeming contradiction arises from the variation in the rate of expansion and contraction with r. A measurement of density is a measurement over a volume and is affected by the differential motion of particles on either side of $r = r_o$. The specified measure of length, however, is a one-dimensional measurement upon a sphere which is contracting and expanding symmetrically.

With this rather fruitful discussion of density, we conclude the presentation of the general theory of relativistic, non-homogeneous models having spherical

symmetry and zero pressure. This general theory can be applied to a variety of problems in which the two basic assumptions of zero (or vanishingly small) pressure and spherical symmetry can be made. Typical problems would be the consideration of low temperature gas or dust clouds, open star clusters, the outer atmosphere of extended stars, clusters of galaxies, or the physical universe itself. In the following and final chapter we shall take the universe for illustrative material. We shall use the general theory which has been presented here and will attempt to construct a non-homogeneous cosmological model which will be consistent with the present observational data.

Chapter V

Application to Cosmology

As an illustration of the foregoing rather esoteric theory, we shall apply it to the entire observable physical universe. The role of this chapter as illustrative, rather than definitive material, must be emphasized. We shall derive several formulae under assumptions which make for simpler mathematics, but which are not the best of physics. However, it will be apparent that the procedures under which these formulae are derived may be generalized to apply to a set of assumptions which will make for better physics. We shall construct a model for the universe which agrees fairly well with the observations. However, it again will be apparent that the constants in this model are adjusted only to the first approximation. By repeated computations it would be possible to adjust these constants more finely so that an even better fit of theory and observation would result. But here again, we shall regard the procedure as being more important than the exact value of numbers in the second and third decimal places.

The cosmological problem not only has a great appeal to the imagination, but it is also, conceptually, another "test" of the General Relativity Theory as a practical law of gravitation. Unfortunately, this is not so in practice. An almost infinite variety of theoretical models may be constructed which will satisfy the Field

Equations of General Relativity. The observational material is a statistical collection of data to a rather limited depth in space. The shortness of the observational range and the dispersion within the data prevent a unique determination of a single theoretical model as being the true one.

The observational data is mainly the work of Hubble (Refs. 15, 16, 17, 18, & 19). It consists primarily of two bodies of information. The first is a statistical correlation between observed red-shift and observed magnitudes. The second is a statistical correlation between observed nebular counts as reduced to certain standard conditions and observed limiting magnitudes. In addition to this moderately well determined data there are two other poorly determined data which must be fitted into our model. The third datum is the present averaged-out density of matter within our own immediate neighborhood of the universe. Not even the proper order of magnitude is known for this important datum. The fourth datum is the "age of the universe", by which we mean the total lapsed time from the last singularity to the present epoch. While there is fair agreement as to the order of magnitude for this datum, there still exists considerable difference of opinion as to its value. It should be remembered that not even the first two bodies of data are sufficiently exact

for our purposes. The really crucial data were obtained at the extreme working range of the Hooker telescope. The empirical corrections which must be applied to this data to reduce it to standard conditions, particularly in the case of the nebular counts, are larger than the effect which is sought. This is hardly a desirable statistical condition, but we shall assume that the data as corrected by Hubble are acceptable. Otherwise we shall have nothing to work with. It is hoped that the use of larger telescopes within the near future will not only extend the data but will also reduce the uncertainties.

A theoretical cosmological model is constructed by making a group of completely arbitrary assumptions as to the distribution of matter within the model, the probable equation of state within the model, the value of the cosmological constant, etc. The consequences of these a priori postulates are then developed according to the General Theory of Relativity and are found as a line-element and equations for proper local density and pressure. If the relationship between the coordinate system and the mass points that populate the model is either known or postulated, these consequences may be further elaborated into expressions for red-shift, total counts to a given coordinate, and observed magnitude for a source of known luminosity at a known coordinate, as seen by an observer

at a specified point within the model. The a priori postulates are so chosen as to simplify the mathematics as much as possible and then their validity is determined by the relative agreement or lack of agreement between the theoretical predictions and the observations.

The simplest possible group of assumptions lead to the highly restricted family of static homogeneous universes. This, of course, is the first logical approximation to the cosmological problem. However, as is well known, the twin observations of a finite density of matter and a real red-shift within the physical universe completely invalidate these simple assumptions.

A more sophisticated set of assumptions, so chosen as to make the partial differential equations linear, yield the more general family of non-static homogeneous universes. Hubble (Ref. 15) has rather prematurely rejected these models as valid representations of the physical universe. Hubble's rejection was based upon three presumptive anomalies for these models. The averaged-out density that was seemingly required was of a higher order than was comparable with other data. A small closed universe whose extent was only slightly greater than the assumed depth of the deepest survey apparently had to be postulated. An age of the universe was indicated which was considerably less than the current estimations of the geological age of the

Earth's crust.

Nevertheless, Hubble's rejection of the homogeneous models is premature for at least two reasons. First, as discussed earlier, the data are not sufficiently good. A slightly different method of handling the corrections would lead to completely different conclusions. Secondly, it is doubtful whether his extrapolation of the red-shift beyond the region in which it has been measured may be accepted. Statistically, the procedure that he followed was the only possible one. The observational data are scanty, but are sufficient to give the coefficient of the first power term to three figures, and to indicate the magnitude and possible range of the coefficient of the second power term, but they are not sufficient to give any information about the third power term. The third power term could quite possibly be insignificant over the range within which the red-shift has been measured, but could be crucially important for the distances of the deepest nebular count surveys. It is obvious that the value of the third power term cannot be obtained from the present observational data, but it could be theoretically predicted for a given model. If a homogeneous model is constructed with the proper values of density, pressure, and cosmological constant to give the first two powers of the red-shift correctly over the measured range, then the

theoretically probable value of the third power term could be calculated from the model. If this third power term were negative and of sufficient magnitude, the homogeneous model could also be fitted to the count data. The usefulness of the model would then depend upon the physical reasonableness of the predicted density, pressure, cosmological constant, and age for the universe. This is an important project which should be carried through, but it is hardly our province since we are concerned with non-homogeneous models.

In order to explain his observational data, Hubble (Ref. 15) has proposed a non-relativistic model which is essentially static and homogeneous but one in which the red-shift is directly proportional to the time that the photon has been in transit between the source and the observer. This makes for statistical simplicity but for serious theoretical complexity. If this proposal were true, problems would arise as to the meaning of measurements made upon extra-terrestrial systems, the known instability of static homogeneous distributions, and the nature of the physical effect which causes the loss of energy in the photon while it is moving through space. Probably the best evidence against this explanation of red-shift has been found by Hubble (Ref. 20). If the red-shift is the linear function of time in flight for the

photon as proposed by Hubble, red-shifts should be found for the galaxies which make up our local system. The more remote members of our local system are at a sufficiently great distance from us so that the expected red-shift should be greater than the corrections which must be applied to the data. However, if the red-shift is due simply to the classic Doppler effect, as postulated in the relativistic models, it should not be found within our own local system since the cluster of nebulae is presumeably gravitationally stable. Hubble's careful measurements of the red-shifts for the members of our local system, after correction for the velocity of the observer, support the latter conclusion. Thus the simplest hypothesis that can be made about the nature of the red-shift is that it is a Doppler effect, produced by actual relative velocities between the nebulae. We must seek a more elevated family of relativistic models which will explain the observational data in the event of the failure of the non-static homogeneous program which previously has been suggested.

In the preceding four chapters we have been concerned with a new set of a priori assumptions which are a higher approximation to the cosmological problem since they allow for non-homogeneity within the models. Tolman (Ref. 2, pp. 174-5) has shown that the homogeneous

models are unstable, going over into non-homogeneous models upon perturbation. It is difficult to conceive of a model which expands from a singular state and which remains homogeneous throughout its history. We have achieved our non-homogeneity at the price of non-linearity in our equations. To make the problem solvable, we have made certain special assumptions which must be recognized because of their bearing upon the cosmological problem.

First, we have assumed zero pressure within our models. This appears to be a good approximation for the present epoch since the important forces presently acting upon the nebulae are probably those of gravitation alone. This is a poor approximation, however, if we wish to extrapolate back in time to the original singular state. A universe composed of mass points without pressure begins its expansion from the singular state at an infinite rate. A real universe, composed of mass volumes and subject to an equation of state, begins its expansion at a finite rate and follows a completely different initial history because of the viscosity-like forces which then operate between the mass volumes. Any calculations of the age of the universe made upon the zero pressure approximation will therefore be in error by being too small.

The error in calculating the age of the universe from models in which pressure terms have been neglected

depends upon the present level of expansion. If the model is well expanded at the present epoch so that the initial stage during which the pressure was important is but a very small fraction of the total history of the model, then the error is slight. However, the model that we shall construct is but partially expanded at present epoch and the pressure was important during a large fraction of its history. Thus we must expect a sizeable error in our calculation of the age of the universe. It should be possible to evolve a correction for the retarding effects of the neglected pressure. Wyman (Ref. 21) has solved the Field Equations for non-homogeneous distributions under an equation of state with certain assumptions. Wyman's solutions break down for the present epoch in which the pressure is negligibly small, whereas our solutions break down for the initial epoch in which pressures are comparable to gravitational forces. A judicious combination of the two solutions should suggest a reasonable correction term. It should be noted, however, that Wyman's solutions presuppose a uniform pressure in all directions at a given point and a given time. We would expect instead, that the pressure at the initial epoch consists of a non-directional chaotic term plus a directed term produced by the outflowing of radiation. Thus a calculation of a correction to the age of the universe would involve several questions of great delicacy.

Secondly, we have assumed a spherically symmetric distribution of matter. This would seem to be a reasonable postulate. The model that we shall use expands from a singular point. The assumption of spherical symmetry about that singular point during the subsequent expansion should be a good first approximation. It is true that we have assumed in the preceding chapters that if a singular point were present within the model that it would be a unique point. Conceptually, space might contain several such singular points, each of which would expand into a gigantic ensemble of nebulae and clusters of nebulae. Such a conceptual ensemble we shall term a meta-galaxy. But the best interpretation of our present data is that the distance of our deepest survey is small compared to the dimensions of our meta-galaxy and that we are completely immersed within it. Thus the existence or non-existence of other meta-galaxies becomes a meta-physical problem. It has no value in the creation of a theory to explain the present observations. Under these conditions, the assumption of spherical symmetry appears to be good physics.

The assumption that the universe expanded from a single singular point is, of course, Lemaitre's hypothesis. Although it is somewhat repugnant because of its overtones of Genesis, we shall see that it is the simplest hypothesis

that we can make. It is significant that theoreticians in other fields, such as Gamow (Ref. 22), have postulated the same ontology to explain the observed relative distribution of the elements.

Thirdly, we have assumed a co-moving coordinate system. That is, in a certain type of coordinate system we have assumed that the mass points have fixed coordinates throughout the history of the model. This implies that there is at least one coordinate system for which the mass points are not in rotation about the center. This is in direct contradiction to the observed behavior of smaller scale gravitational units such as nebulae and clusters of nebulae. We cannot fit a co-moving coordinate system to the mass points that make up a rotating spiral nebula because the variation of angular velocity with radius vector would destroy the coordinate system within a finite time. We must assume that the meta-galaxy is about the only gravitational system which is not in rotation, but this is not too unreasonable. First, the meta-galaxy is unstable since it is observed to be expanding. Rotation would tend to produce a stable system. Secondly, in the singular state there must be at least one coordinate system for which Lemaitre's "giant atom" is not in rotation. Conservation of angular momentum would then require that the expanded universe be rotationally at rest with respect to

this special system of coordinates. Thus the assumption of co-moving coordinates seems to be a reasonable hypothesis.

In addition to the three assumptions which are inherent in the general theory developed in the preceding chapters we shall make three special assumptions which are frankly designed to reduce the mathematical labor. Two of these special assumptions are reasonable, while the third is actually bad physics. As the fourth assumption we shall take the cosmological constant to be zero. This has the great advantage of giving non-elliptic solutions. As pointed out in the preceding chapter, this is the most reasonable a priori value to assume for this term. But this remains a special assumption which we are not compelled to make. Interesting and useful solutions might well result from assuming small positive or negative values for the cosmological constant. This possibility should be kept in mind for future work.

The fifth assumption is that the observer is is located at the center of the meta-galaxy. This homocentric hypothesis is manifestly poor physics. It can be rationalized as being the logical first step in the transition from homogeneous to non-homogeneous models. But even this rationalization implies that expressions for a non-central observer will be sought immediately after a workable procedure is established. Since the purpose of

this chapter is illustration rather than deduction of exact cosmological conclusions, the homocentric hypothesis is excusable. It does simplify the mathematical work considerably. However, all of the quantities having cosmological interest; such as density, curvature, red-shift, counts, magnitudes, and age, depend upon the choice made for the position of the observer. A complete satisfactory cosmological theory cannot be expected until the more general case of the non-centric observer has been solved. The procedures which will be presented in this chapter appear to be capable of such a generalization.

Under the first four assumptions there are only three possible solutions to the Field Equations. They are M_1 behavior with equations (4.2) and (4.3): O_1 behavior with equations (4.4) and (4.5); and M_2 behavior with equations (4.6) and (4.7). We could construct a model containing any arbitrary combination of these three solutions. For simplicity we shall restrict our meta-galaxy to a single type of behavior. This is a sixth assumption. As a first approximation it is entirely logical.

We can show by appealing to formulae which will be presented later that the M_2 solution does not lead to a satisfactory model. In the neighborhood of the observer at the origin, we must choose a coordinate system so that

Table VII

Model	Density gms./cc.	Parameter	R lt.-yrs.	T yrs.
M_1	10^{-30}	7.616	1.86×10^9	1.85×10^9
	10^{-28}	2.905	2.08×10^9	1.58×10^9
	$= 5.077 \times 10^{-28}$		Transition point	
O_1	10^{-27}	$89^\circ\ 7'$	1.89×10^9	8×10^8

$y \approx r$ to make the coordinate distance and the observer's proper distance agree. If some other coordinate system were used, it would have to be transformed to an equivalent form when calculations were made with the model to be compared with the observations. Since a radial increment and the proper distance must agree in the neighborhood of the origin, it can be shown that $g(r) \approx b_n r^n$ where $n > 0$ in that neighborhood. But the increment in nebular counts in the neighborhood of the origin must be $dN \approx (\text{constant})\ r^2$ from which we must have $dh/dr \approx C_1 r^2$, where $C_1 > 0$ in this region. But since $h(r) < 0$ is one of the basic conditions for the existence of the M_2 solution, we must furthermore have $h(r) \approx - C_2 + C_1 r^2 + -$, where $C_2 > 0$ in this region. But then the parametric equation (4.6) for the M_2 solution gives $y \approx (\text{constant})/r^n \neq r$ in the neighborhood of the origin which is in direct contradiction to the first necessary assumption. Thus the M_2 solution must be rejected.

There remain two possible solutions which may be used for our model. These two solutions are non-overlapping and the choice of one or the other depends entirely upon the proper density at the origin at the present epoch as may be seen in Table VII which has been calculated by methods which are yet to be presented. The second column gives the assumed density at the origin and for the present epoch. The third column gives the value of the parameter to be used in equations (4.2), (4.3), (4.4), and (4.5) under these conditions. The next column gives local radius of curvature while the last column gives present age of the universe.

Since we wish to keep the density within our model as low as possible, we shall chose the M_1 solution. This choice is not as arbitrary as it may seem since it is based upon considerable computation which will not be presented here. All formulae and calculations in the remainder of this chapter will be given in terms of the M_1 solutions. However, calculations have been made with both types of models and in all cases it has been found that a given M_1 relationship can be transformed into its equivalent O_1 relationship by altering a few signs and replacing the hyperbolic functions with the corresponding circular functions.

Within a sufficiently small neighborhood of the origin, the model will be essentially homogeneous. Hence,

within that neighborhood, we must have $g(r) \approx r^2/R^2$ and $h(r) \approx 4\, k^3\, r^3$ by the reasoning of the previous chapter. The constant k is a scale-constant which will be used to adjust the co-moving coordinate system so that it agrees with proper measure at the present epoch. Consequently, the functions $h(r)$ and $g(r)$ may be expressed for the entire model as power series in the form

$$\begin{aligned} h(r) &= 4\, k^3\, r^3\, (1 + a_1 r + a_2 r^2 + a_3 r^3 + - -) \\ g(r) &= (r^2/R^2)\, (1 + b_1 r + b_2 r^2 + \; - \; -), \end{aligned} \tag{5.1}$$

where the coefficients a_i and b_i are determined by the non-homogeneity of the model.

It was also shown in the preceding chapter that we must take $f(r) =$ constant for both the M_1 and the O_1 solutions if negative densities are to be avoided. This conclusion is quite obvious if either equations (4.2) and (4.3) or equations (4.4) and (4.5) are used with equation (1.36). Since we must take $f(r)$ to be a constant, we shall assign it the value of zero. This places the origin of the time coordinate at the singularity. Consequently, we shall write for the time coordinate in equation (4.3) $t + T$, where t is the local time of the observer which reduces to zero at the present epoch, while T is the present age of the model.

Using these conventions, we can rewrite equations (4.2) and (4.3) as

$$y = 2k^3R^2r\,\frac{(1 + a_1r + a_2r^2 + a_3r^3 + \cdots)}{(1 + b_1r + b_2r^2 + \cdots)}\left[\cosh\theta - 1\right] \tag{5.2}$$

$$t + f = 2k^3R^3\,\frac{(1 + a_1r + a_2r^2 + a_3r^3 + \cdots)}{(1 + b_1r + b_2r^2 + \cdots)^{3/2}}\left[\sinh\theta - \theta\right] \tag{5.3}$$

For convenience in calculation, we will wish for our line-element to reduce to the Special Relativity form in the immediate neighborhood of the observer. By equations (4.12) and (4.13) of the preceding chapter, this will require that $y \approx r$ in that region. This, in turn, places the condition on (5.2) that

$$2k^3R^2(\cosh\theta_o - 1) = 1 \tag{5.4}$$

where θ_o is the value of the parameter for the origin and the present epoch.

When equation (5.4) is satisfied, we may develop $y(r,t)$ over the entire model as a double Taylor expansion in the form

$$y = r\left[1 + c_1t + c_1^or + c_2t^2 + c_{12}rt + c_2^or^2 + c_3t^3 + c_{13}rt^2 + c_{23}r^2t + c_3^or^3 + \cdots\right] \tag{5.5}$$

We shall furthermore place a condition on the series coefficients which bear superscripts so that no term in r alone to a higher power than unity will appear in (5.5).

That is, we shall make

$$c_1^o = c_2^o = c_3^o = - - - = 0, \tag{5.6}$$

This is actually a condition upon the a_i and b_i coefficients of (5.1). In practice, we shall determine the a_i coefficients by the cosmological requirements of the model and then determine the b_i coefficients so that condition (5.6) is satisfied. This condition makes it possible to at least approximately integrate the equations for the geodesics. Moreover, it puts (5.5) into an analogous form with the equivalent expression for a homogeneous model. The cross-terms, whose coefficients are c_{12}, c_{13}, c_{23}, etc. produce the non-homogeneity in the model. If these cross-terms should vanish and if condition (5.6) is satisfied, then $y(r,t) = r\,F(t)$, which is exactly the homogeneous form. We should note that we now have only one arbitrary function at our disposal. The remaining arbitrary function is $h(r)$ since $g(r)$ will now be determined by $h(r)$ and condition (5.6)

The desired relation for nebular counts is not too difficult to derive. By combining the line-element (1.23) with the density equation (1.31) and integrating over the angle variables, we see that the amount of mass between r and $r + dr$ is

$$dM = \frac{\frac{dh}{dr}}{2\sqrt{1+g}}\, dr \tag{5.7}$$

Expanding this expression into a series and integrating the series term by term, we find that the total mass out to a general coordinate distance r is

$$M = 2k^3\left[r^3 + a_1 r^4 + \left\{a_2 - 0.3\,\frac{1}{R^2}\right\} r^5 + \left\{a_3 - \frac{1}{R^2}\left(\frac{a_1}{3} + \frac{b_1}{4}\right)\right\} r^6 + - - -\right] \tag{5.8}$$

The mass, of course, is independent of the time since co-moving coordinates have been used. In the neighborhood of the origin, we must have $2\,k^3 r^3 = \frac{4}{3}\pi \rho_o\, r^3$ which determines the scale-constant k as

$$k^3 = \frac{2}{3}\pi \rho_o \tag{5.9}$$

where ρ_o is the proper density at the origin and at the present epoch.

The total number of nebulae out to a given coordinate distance r will be proportional to (5.8). We can parallel the reasoning of Tolman and Hubble (Refs. 23 and 15) for the counts within a homogeneous model exactly in all details and write to the same order of approximation for the counts within our non-homogeneous model

$$\text{Log } N = 0.6\,(m - 4\frac{\delta\lambda}{\lambda}) + F^* + C, \tag{5.10}$$

where N is the total number of nebulae counted per unit area of the sky down to a limiting magnitude m. This

limiting magnitude must be corrected for the change in the rate of energy flow produced by the red-shift and for the different absorption in the Earth's atmosphere and in the telescope mirror, and the different sensitivity of the photographic plate for the shifted radiation. The $4\frac{\varepsilon\lambda}{\lambda}$ term is Hubble's approximation for this total red-shift correction. The red-shift to be used is that for a nebulae having an absolute magnitude $\overline{M}$, taken by Hubble to be -15.15, and having an apparent magnitude equal to the limiting value m of the counts. There is some question as to whether this approximation for the total red-shift correction is valid, but for uniformity we shall use it. The constant C depends upon the conventions that are adopted. Hubble reduced his counts to nebulae per square degree and found statistically that $C = -9.052$. The term F^* contains the effects of the non-homogeneity and the curvature of the model. By making the parallel derivation referred to, we find that

$$F^* = \log\left[1 + a_1 r + \left\{a_2 - 0.3\frac{1}{R^2}\right\}r^2 + \left\{a_3 - \frac{1}{R^2}\left(\frac{a_1}{3} + \frac{b_1}{4}\right)\right\}r^3 + \cdots\right] \quad (5.11)$$

In this derivation we have used a relation between coordinate distance and limiting magnitude which is yet to be given.

The derivation of the red-shift is not quite so simple. It can be shown that the radial lines of our coordinate system are null-geodesics. Then light, leaving

a source at r_1 and traveling down a radial line to the origin, must by the line-element be subject to the differential relation

$$-\frac{1}{\sqrt{1+g}}\frac{\partial y}{\partial r}\,dr = dt \tag{5.12}$$

By using (5.5) and (5.1) this differential relation may be expressed as an infinite series in the form

$$-\left[1+c_1t+c_2t^2+2c_{12}rt-\frac{r^2}{2R^2}+c_3t^3+2c_{13}rt^2+3c_{23}r^2t-\frac{c_1}{2R^2}r^2t-\frac{b_1}{2R^2}r^3+\cdots\right]dr = dt \tag{5.13}$$

If we consider a ray of light leaving a source at the coordinate r_1 at the time t_1 and arriving at the origin at the time t_o which is taken to be very small, being comparable to zero, we will have formally that

$$\int_o^{r_1}\left[\text{——}\right]dr = t_o - t_1 \tag{5.14}$$

Where the integrand is the infinite series of (5.13). Unfortunately the integrand contains terms in both r and t in such a form that they cannot be separated. However, when the photon passes through any given value of r, it does have a definite value of t. If we knew this relation between r and t along the radial line, we could integrate (5.14) directly.

We can approach this desired relationship between r and t by successive approximations. If we take only the

first term of the infinite series, we find the first approximation to be

$$t = t_o - r.$$

If we substitute this first approximation back into the integrand of (5.14), now taking the first two terms, we find the second approximation to be

$$t = t_o - (1 + c_1 t_o)r + \tfrac{1}{2}c_1 r^2$$

Substituting this second approximation back into the integrand, now taking all terms up to the second power, we shall find the third approximation. Substituting the third approximation back into the integrand and using more terms will give a still higher order of approximation. But, by this time, the approximation has become very awkward to handle.

However, we are interested only in obtaining the red-shift. For this let us consider two light rays; one leaving r_1 at t_1 and arriving at the origin at $t_o = 0$, while the other leaves r_1 at $t_1 + \delta t_1$ and arrives at the origin at $t_o = \delta t_o$. By making the fourth approximation on the two light rays simultaneously and subtracting them, we see that large blocks of terms will be cancelled. Furthermore, by taking t_o to be so small that all of its higher powers greater than unity may be neglected, we shall find that

$$\frac{\delta t_1}{\delta t_o} = 1 - c_1 r_1 + \left\{\frac{1}{2}c_1^2 + c_2 - c_{12}\right\} r_1^2 - \left\{\frac{1}{6}c_1\left(c_1^2 - \frac{1}{R^2}\right) + \frac{4}{3}c_1 c_2 - c_1 c_{12} + c_3 - \frac{4}{3}c_{13} + c_{23}\right\} r_1^3 + \cdots \quad (5.15)$$

But since $\frac{\delta t_o}{\delta t_1} = 1 + \frac{\delta \lambda}{\lambda}$, we can solve (5.15) for the red-shift. Thus the red-shift, observed at the origin and the present epoch for a source which is located at r, will be

$$\frac{\delta \lambda}{\lambda} = c_1 r + \left\{\frac{1}{2}c_1^2 - c_2 + c_{12}\right\} r^2 + \left\{\frac{1}{6}c_1\left(c_1^2 - \frac{1}{R^2}\right) - \frac{2}{3}c_1 c_2 + c_1 c_{12} + c_3 - \frac{4}{3}c_{13} + c_{23}\right\} r^3 + \cdots \quad (5.16)$$

This relation can be checked in two ways. Our model will reduce to a homogeneous one if the cross-terms vanish, that is, if $c_{12} = c_{13} = c_{23} = 0$. Then the red-shift should also reduce to the homogenous relation under this condition. If the well known equation for red-shift in a homogeneous model is re-expressed in equivalent form, this will be seen to be true. The second check may be effected by letting all terms in t^2 and higher powers of t vanish. Then the instantaneous acceleration of the source with respect to the origin also vanishes and the Special Relativity equation for Doppler shift can be used. This limiting case is also verified. Thus equation (5.16) would seem to be a valid expression for the red-shift in our non-homogeneous model.

The last relation that must be developed is one between coordinate distance and observed magnitude. We first shall be concerned with conceptual quantities which are called bolometric. These are the quantities which would

conceptually be measured upon a perfect photographic plate exposed in a perfect telescope which is located in empty space. Bolometric quantities may be reduced to the photographic quantities which are actually measured by means of a series of semi-empirical corrections which will allow for the selective absorption through the Earth's atmosphere, the imperfections of the telescope, and the spectral sensitivity of the photgraphic plate.

A magnitude measurement is based upon the observed density of a photographic image which was exposed for a given time. Thus we are concerned with the amount of energy which flows through our telescope aperture in a unit time. Let us take the center of our telescope aperture to be exactly at the coordinate origin. Every photon that leaves the source at coordinate r_1 travels along a null-geodesic line in our co-moving coordinate system. We are concerned with the very minute cone of null-geodesics which has its apex at the source and which just fills the telescope aperture. Energy flows into the cone from the source at a rate which we shall denote as $\Delta E_1 / \Delta t_1$ but flows out of the cone into the telescope at a different rate, which we shall denote as $\Delta E_o / \Delta t_o$. The difference in the entrance and exit rates of energy flow is caused by the relative motion of the source and the telescope. Since our model is expanding, each photon that enters our telescope

is degraded in energy by the red-shift. Moreover, a block of energy which flows into the cone from the source in unit time will require more than unit time to flow into the telescope. Thus we have two effects which will reduce the rate of energy flow into our telescope from the original rate of energy flow into the cone from the source. We have, in fact

$$\Delta E_0 = \left(\frac{\lambda}{\lambda+\delta\lambda}\right)\Delta E_1 \qquad \Delta t_0 = \left(\frac{\lambda+\delta\lambda}{\lambda}\right)\Delta t_1 \qquad \frac{\Delta E_0}{\Delta t_0} = \left(\frac{\lambda}{\lambda+\delta\lambda}\right)^2 \frac{\Delta E_1}{\Delta t_1} \tag{5.17}$$

The central line of the cone of null-geodesics will be a radial line of the coordinate system. This radial line is also one of the null-geodesics. Let us consider the small angle ω measured in proper units at the source between this radial line and one of the extreme rays of the cone. Then the rate at which energy flows into the cone from the source will be

$$\frac{\Delta E_1}{\Delta t_1} = (\text{Constant})\, I\, \omega^2 \tag{5.18}$$

where I is the total emissive power of the source and the constant contains the various factors of proportionality. If we can determine the angle ω, we can solve our problem.

We shall solve the problem by determining the null-geodesics through the source. We shall find to a very close approximation the null-geodesic which just barely

enters the aperture of our telescope. We may then find the angle ω' as measured in coordinate units at the source between the radial line and this extreme null-geodesic by

$$\tan \omega' = -r \frac{d\theta}{dr} \tag{5.19}$$

as is well known in analytic geometry. But the angle ω' measured in coordinate units cannot be used in (5.18). The relationship between coordinate measure and proper measure is given, of course, by the line-element (1.23). We can see from this line-element that the angle ω measured in proper units at the source must be

$$\tan \omega = -\frac{y\sqrt{1+g}}{y'} \frac{d\phi}{dr} \tag{5.20}$$

from which the problem can be solved by combining (5.20), (5.18), and (5.17).

The geodesic equations are found without difficulty by the use of Dingle's formulae (ref. 1, pp. 254-7). They are

$$\frac{d^2r}{ds^2} + \left[\frac{y''}{y'} - \frac{\frac{dg}{dr}}{2(1+g)}\right]\left(\frac{dr}{ds}\right)^2 + 2\frac{\dot{y}'}{y'}\frac{dr}{ds}\frac{dt}{ds} - \frac{(1+g)y''}{y'}\left(\frac{d\theta}{ds}\right)^2 - \frac{(1+g)y}{y'}\sin^2\theta\left(\frac{d\phi}{ds}\right)^2 = 0 \tag{5.21}$$

$$\frac{d^2\theta}{ds^2} + 2\frac{y'}{y}\frac{dr}{ds}\frac{d\theta}{ds} + 2\frac{\dot{y}}{y}\frac{d\theta}{ds}\frac{dt}{ds} - \sin\theta\cos\theta\left(\frac{d\phi}{ds}\right)^2 = 0 \tag{5.22}$$

$$\frac{d^2\phi}{ds^2} + 2\frac{y'}{y}\frac{dr}{ds}\frac{d\phi}{ds} + 2\cot\theta\frac{d\theta}{ds}\frac{d\phi}{ds} + 2\frac{\dot{y}}{y}\frac{d\phi}{ds}\frac{dt}{ds} = 0 \tag{5.23}$$

$$\frac{d^2t}{ds^2} + \frac{y'\dot{y}'}{1+g}\left(\frac{dr}{ds}\right)^2 + y\dot{y}\left(\frac{d\theta}{ds}\right)^2 + y\dot{y}\sin^2\theta\left(\frac{d\phi}{ds}\right)^2 = 0 \qquad (5.24)$$

The notation is one which was previously used in which the dots represent partial differentiation with respect to t while the primes represent partial differentiation with respect to r.

If in (5.22) we take $\theta = \frac{1}{2}\pi$ and $\frac{d\theta}{ds} = 0$ as initial conditions, we see that $\frac{d^2\theta}{ds^2} = 0$. This means that if a geodesic initially lies within this plane, it lies within the same plane throughout its trajectory. In the same way in (5.23) if the geodesic was initially a radial line with $\frac{d\phi}{ds} = 0$, then $\frac{d^2\phi}{ds^2} = 0$ and the geodesic remains a radial line. This verifies the earlier statements that the radial lines are null-geodesics. But only one geodesic within the cone will be a radial line. For all of the others $\frac{d\phi}{ds}$ will have a non-zero initial value. However, with $\theta = \frac{1}{2}\pi$ we see that (5.23) may be integrated exactly as

$$\frac{d\phi}{ds} = \frac{K_1}{y^2} \qquad (5.25)$$

where K_1 is a constant of integration.

By using the line-element in a degenerate form since $ds = 0$ for a null-geodesic, and by using the initial condition $\theta = \frac{1}{2}\pi$ we may restate equation (5.24) in the form

$$\frac{d}{dt}\log\left(\frac{dt}{ds}\right) = -\frac{\dot{y}'}{y'} + \left[y^2\frac{\dot{y}'}{y'} - y\dot{y}\right]\left(\frac{d\phi}{dt}\right)^2 \qquad (5.26)$$

In the homogeneous special case in which the cross-terms vanish we can integrate (5.26) exactly as $\frac{dt}{ds} = K_2 r/y$ where K_2 is another constant of integration. Then, by using this result along with (5.25) in the line-element we can integrate exactly and find the homogeneous null-geodesics. We would find that the homogeneous null-geodesics in the closed O_1 models would be ellipses centered on $r = 0$ and having semi-major axes of R and semi-minor axes which depend on the initial angle of of the ray at the source. We would be concerned, of course, with that extreme null-geodesic whose semi-minor axis is equal to the radius of our telescope aperture. Similarly, in the open M_1 models we would find the homogeneous null-geodesics to be hyperbolas centered on $r = 0$ and having semi-major axes of R and semi-minor axes which again depend upon the initial angle of the ray at the source.

We cannot integrate (5.26) exactly for our non-homogeneous models; hence we must make an approximate solution. We shall assume that our non-homogeneous model is nearly homogeneous. Accordingly, the cross-terms in (5.5) will be vanishingly small as compared to the purely time dependent terms. We will take as our approximate expression for $y(r,t)$ the purely time dependent

terms plus a small non-homogeneous correction term which will be the largest of the cross-terms. For convenience in notation, we shall denote $K_1/K_2 = \delta$. Then to the desired order of approximation we find that $d\phi/dt \approx \delta/ry$. Using this result and the approximate expression for $y(r,t)$ in (5.26) we find that

$$\frac{dt}{ds} \approx K_2 \frac{r}{y} \frac{(1+\frac{1}{3}c_1c_{12}r^3+---)}{r^{c_{12}}\delta^2} \tag{5.27}$$

Using (5.27) and (5.25) in the line-element, we find that the differential equation which must be satisfied by the null-geodesics in our co-moving coordinates is

$$\frac{y'^2}{y^2(1+g)}\left(\frac{dr}{d\phi}\right)^2 \approx \frac{r^2}{\delta^2}\frac{(1+\frac{1}{3}c_1c_{12}r^3+---)^2}{r^{2c_{12}}\delta^2} - 1 \tag{5.28}$$

The solution of this differential equation is a distorted conic section. The semi-minor axis of this distorted conic section is determined by the condition $dr/d\phi = 0$. Because of the extreme smallness of the remaining factors on the right-hand side of (5.28), this condition is satisfied to a very high order of approximation by $r = \delta$. Thus, to obtain the initial angle of the extreme ray within the null-geodesic cone, we shall take δ to be the radius of the aperture of our telescope. Then we shall have

$$\omega \approx \tan\omega \approx \frac{\delta}{r}\frac{(1+c_{12}\delta^2\log r+---)}{(1+\frac{1}{3}c_1c_{12}r^3+---)} \tag{5.29}$$

By combining (5.29), (5.18), and (5.17) we find that

$$l_b \approx \frac{\text{Constant}}{r^2}\left(\frac{\lambda}{\lambda+\delta\lambda}\right)^2\left[\frac{1+C_{12}\delta^2\log r+---}{1+\frac{1}{3}C_1C_{12}r^3+---}\right]^2 \qquad (5.30)$$

where l_b is the bolometric luminosity of the source as observed at the origin. For the deepest count in the model that we shall use, the non-homogeneous correction term in (5.30) has a value of 0.9966, which differs insignificantly from unity for all observational purposes. Thus while a finite focusing effect does exist because of the non-homogeneity of the model, it has no practical consequences for a central observer. This is not to say, however, that the focusing effect would not be important for a non-central observer. In the latter case, the term δ would be of astronomical instead of terrestrial dimensions.

Since the non-homogeneous focusing effect is insignificant for the central observer within our model, we shall henceforth neglect it. Following the reasoning of Tolman and Hubble (Ref. 23) we can write for the desired relation between the coordinate distance and the observed photographic magnitude for a source of absolute photographic magnitude M located at r and having an observed red-shift of $\frac{\delta\lambda}{\lambda}$ as

$$\log r = 0.2\left\{m-5\log\left(1+\frac{\delta\lambda}{\lambda}\right)-K-M\right\}+1 \qquad (5.31)$$

where the term K consolidates the semi-empirical correction terms which are necessary to change from conceptual bolometric to actual photographic magnitudes. Hubble (Ref. 15) has stated that to a close approximation $5 \log(1 + \frac{\delta\lambda}{\lambda}) + K = 4 \frac{\delta\lambda}{\lambda}$; consequently we shall use (5.31) in the form

$$\log r = 0.2(m - 4 \frac{\delta\lambda}{\lambda}) + 4.543 \qquad (5.32)$$

for making computations, where we have taken $M = -15.15$.

Equations (5.32), (5.16), and (5.10) are the required relations for making computations with our model. As stated earlier, the relation (5.32) was used in the derivation of (5.10). In making computations with these formulae we must use a consistent set of relativistic units. The particular set of units that we shall use in calculating from our model will be the year as the unit of time, the light-year as the unit of length, and a comparable unit for mass.

We shall fit our model to a first approximation to the primary body of data, namely the red-shift against magnitude and the counts against limiting magnitude. The remaining two pieces of data, namely the local density and the age of the universe, will be determined by the requirements of the fitting procedure for the previous data. In all cases we shall accept Hubble's figures as published.

Since these were collected at the working limits of the largest telescopes and were corrected by factors which lie on the border-line of knowledge, there is reasonable doubt about the exact values of the figures and the correctness of the procedures. Among these doubtful items are the determination of the limiting magnitudes in the counts, the value of the term K which corrects from bolometric to photographic magnitudes, the approximating procedures, the calculation of the term Δm_o, and the correct value of $\overline{M}$. For instance, Fletcher (Ref. 24) has recently suggested that $\overline{M}$ should be about a half magnitude fainter than the value that was used.

We do not wish to begin a controversy. Hubble's figures and data may be exactly correct in all particulars. We shall accept his published figures and shall attempt to construct a model, to the first approximation only, based upon presently accepted physics which will justify Hubble's data. This is an important desideratum since Hubble's work has been widely interpreted as demostrating a kind of failure of the classical relativity theory and of requiring the discovery of a new principle of physics. However, because of reasonable doubt as to the complete validity of the experimental data, the fault for any partial failure of this attempt must not be assessed entirely against the theory.

Hubble has derived a statistical relation between red-shift and relativistic coordinates as

$$\frac{\delta\lambda}{\lambda} = 5.37 \times 10^{-10}\, r + 2.54 \times 10^{-19}\, r^2. \qquad (5.33)$$

The first term is determined to three figures from the data and can be considered to be quite good. The data are too scant to fix the second term. Hubble states that the probable value for the second term must lie between 2.26×10^{-19} and 3.2×10^{-19}. Because of the lack of sufficient data, it is impossible to determine anything about a possible third power term.

We shall take a model in which the third power term of the red-shift is negative and important in value for the deeper count surveys. Accordingly, Hubble's statistical adjustment of the second power term would be a bit too low. We shall arbitrarily take his upper probable limit as being the correct value. This is quite crude, but is possibly sufficient for a first approximation. A correct procedure would be to calculate the red-shift against magnitude relation for the model and then to make the model fit the observations for the cluster nebulae. This latter procedure will be required for a more detailed fit of model against observations.

Hubble has statistically determined three terms of (5.10). The observationally determined terms are

Log N, m, and C. It is conceptually possible to observationally determine $\frac{\delta\lambda}{\lambda}$, but practically this term was out of range of existing instruments. Therefore, there are two terms, namely $\frac{\delta\lambda}{\lambda}$ and F^*, which must be theoretically determined from the model to fit the observations. Hubble has defined a quantity Δm_o as

$$\Delta m_o = m - \frac{10}{6}[\log N - C] = 4\frac{\delta\lambda}{\lambda} - \frac{10}{6}F^x \tag{5.34}$$

which can be evaluated from the observational data. Fitting our model to the count data then reduces to obtaining the proper values of Δm_o. Hubble has tabulated his values of Δm_o against limiting magnitudes. However, if we calculate this quantity from his published data, we will find slightly different values as shown in Table VIII. As before, we shall accept Hubble's values in making our calculations. The differences shown in the last column of the table require explanation, however.

Turning finally to our theoretical model, the conditions (5.4) and (5.9) when solved with the expression for c_1 gives

$$\cosh\theta_o + 1 = c_1^2/2k^3 = 3c_1^2/4\pi\rho_o \tag{5.34}$$

Since the numerical value of c_1 is given in (5.33), the value of the parameter at $r = t = 0$ can be calculated whenever ρ_o is assumed. This value of the parameter when

Table VIII

m	Hubble	Omer	H - O
21.03	0.676	0.673	+ 0.003
20.0	0.468	0.44	+ 0.028
19.4	0.368	0.41	- 0.042
19.0	0.314	0.31	+ 0.004
18.47	0.253	0.24	+ 0.013

used with the expression for c_1 and equations (5.2) and (5.3) can be used to determine R and T. This was the procedure by which Table VII was calculated.

We shall take a model in which $\rho_o = 3.3 \times 10^{-28}$ gms/c.c. Then from (5.34) we find that $\Theta_o = 1.3603$ and that $R = 3.147 \times 10^9$ lt.-yrs. while $T = 1.344 \times 10^9$ yrs. This assumption of density is not arbitrary. This is the result of a series of calculations in which various different densities were assumed until a negative third power term was found for the red-shift as given by (5.16) which was thought to be of the proper order of magnitude.

The next step is to fit to the second power term of the red-shift. Of the three elements that make up this term in (5.16), $\frac{1}{2}c_1^2$ is fixed, c_2 is determined by the assumption for density and is a maximum for the transition

density between the O_1 and M_1 models, while c_{12} is determined by the non-homogeneity of the model. If the model were homogeneous, the c_{12} term would vanish. By the extreme assumption that the density of a homogeneous model were exactly the transition value, we would have $\frac{1}{2}c_1^2 - c_2 \approx 2.16 \times 10^{-19}$, which is rather close to Hubble's probable lower limit for this term. However, such a homogeneous model would not yield Hubble's count data. To fit a homogeneous model we would have to prove either that Hubble's count data was very seriously in error, or we would have to construct a different homogeneous model having non-zero pressure and cosmological constant.

Under the assumption of zero cosmological constant we must assume a non-homogeneous model. For the model that we are constructing we will have $\frac{1}{2}c_1^2 - c_2 = 1.916 \times 10^{-19}$. Since we wish the coefficient of the second power term to be 3.2×10^{-19}, we must take $c_{12} = 1.284 \times 10^{-19}$. The condition that $c_1^0 = 0$ gives $b_1 = -3.247\, a_1$. This result and the preceding numerical value for c_{12} requires that $a_1 = -9.8 \times 10^{-10}$. Thus, to the first approximation, we are required to assume a non-homogeneous model in which the density is decreasing away from the center.

The third step is to satisfy the count data. We shall assume in the shallow survey to the limiting magnitude of 18.47 that the effects of the third power term in the red-shift may be ignored. This a poor assumption but it is justified for a first approximation. Then, to obtain Hubble's value of Δm_0 we must take $a_2 = 1.234 \times 10^{-17}$. The condition that $c_2^0 = 0$ then gives $b_2 = -4.015 \times 10^{-17}$. We may now calculate the third power term of the red shift. The most important element in this term as given by (5.16) is c_{23}. We find that $c_{23} = -1.642 \times 10^{-27}$ under the foregoing assumptions. Calculating the remaining elements of the third power term for the red shift we find that for this model

$$\frac{\delta\lambda}{\lambda} = 5.37\times10^{-10} r + 3.2\times10^{-19} r^2 - 1.583\times10^{-27} r^3 + \cdots \qquad (5.35)$$

in which the third power term is of the proper order of magnitude but is probably too great in value.

The last step is to fit the remainder of the count data. We still have a_3 at our disposal and we shall attempt to tailor it to give fair agreement with the data. In making this fit we shall assume that the effects of the fourth power term in the red-shift may be neglected throughout the range in which count data are available. For the deepest count to a limiting magnitude of 21.03 this is probably a poor assumption. Again, for a first

approximation the procedure is justified.

To make the required calculations we must find the values of r and $\frac{\delta\lambda}{\lambda}$ which correspond to the limiting magnitudes m. To a first approximation we can do this graphically by plotting two curves to a sufficiently open scale so that three figure interpolation is possible. We have plotted (5.35) to such a scale where we have taken convenient values of r. We have then used these data in (5.32) which was rewritten as

$$m = 5 \log r - 22.715 + 4 \frac{\delta\lambda}{\lambda} \tag{5.36}$$

and have plotted a second curve of limiting magnitude against r. From these two curves the required values of r and $\frac{\delta\lambda}{\lambda}$ for the limiting magnitudes m were found and they are given in the first three columns of Table IX. This information along with Hubble's values for Δm_o given in Table VIII determine the values of the relativistic correction term F^* which will be required to justify the count data. This partially experimental quantity, which has been denoted as F_o^*, is listed in the fourth column of Table IX. The theoretic value of F^* as determined by our model is given by (5.11). Since the coefficients a_1, a_2, b_1, and R have already been fixed by the requirements upon our model, we have only a_3 to vary to fix the data. This was done by computing those terms of (5.11) which are fixed and

subtracting then from $\log^{-1} F_o^*$. The differences should be mainly the $a_3 r^3$ term, although the neglected higher power terms in (5.11) might be important.

Table IX

m	$\frac{\delta\lambda}{\lambda}$	$r \times 10^{-8}$	F_o^*	F_c^*	C - O	$\frac{\rho}{\rho_o}$
21.03	0.165	4.19	- 0.0096	- 0.0264	- 0.0168	0.43
20.00	0.138	2.72	+ 0.0504	+ 0.0603	+ 0.0099	1.18
19.4	0.110	2.16	+ 0.0432	+ 0.0473	+ 0.0041	1.18
19.0	0.095	1.85	+ 0.0396	+ 0.0354	- 0.0042	1.15
18.47	0.082	1.49	+ 0.0450	+ 0.0199	- 0.0251	1.10

As would be expected from the approximate procedure used in fitting this model, this predicted $a_3 r^3$ term nearly vanishes for the survey to the limiting magnitude 18.47. The remaining four surveys were used to determine an average value for a_3 which is $a_3 = -2.46 \times 10^{-26}$. Using this average value in the full form of (5.11) the fifth column of calculated values of F^* was computed. The residuals, which are the differences between the calculated and observed values of F^*, are given in the sixth column. The agreement between this first approximate form of the model and the observations is fair. The residuals for the three intermediate surveys are essentially equal to the residuals obtained by Hubble (Ref 15, Table X) from

his static model with a unknown cause for the red-shift. The congruity of the two extreme surveys is disturbed by the known errors in this first approximating procedure, which have been discussed earlier. This agreement can obviously be improved by repeating the indicated procedures several times with increasingly better approximations at each calculational step.

We may compute the local proper density anywhere within the model from the use of (1.36). For the particular time $t = 0$, this density expressed as a ratio to the proper density at the origin is

$$\frac{\rho}{\rho_0} = 1 + \frac{4}{3} a_1 r + \frac{5}{3} a_2 r^2 + 2 a_3 r^3 + - - - \qquad (5.37)$$

Since the coefficients a_1, a_2, and a_3 are now fixed, this density ratio may be computed, where again the effects of powers higher than the third have been neglected. This information is tabulated in the last column of Table IX. The density ratio slowly decreases away from the origin to a minimum of about 0.99 and then increases to a maximum of 1.18 within the range of the count surveys, but starts to drop drastically at the extreme range of the surveys. This is the type of behavior that we would expect, although the model needs further adjusting. If the singular state of the universe had a finite mass, we would expect the density ratio to begin to drop when the coordinate distance

became comparable to the age of the universe. However, from present estimations as to the age of the universe, we would expect the density ratio to begin to drop for coordinates about three times greater than the one indicated here. There are at least two causes for the seeming anomaly. First, the coefficient of the third power term in the red-shift given in (5.35) is probably too great. Second, higher power terms have been consistently neglected throughout the computations. These neglected higher power terms might well be significant for the deepest surveys.

We can summarize this first tentative model by listing the coefficients which have been employed. They are

$\rho_o = 3.3 \times 10^{-28}$ gm/cc. $\quad R = 3.147 \times 10^9$ lt-yrs.

$\Theta_o = 1.3603 \quad T = 1.344 \times 10^9$ yrs.

$a_1 = -9.8 \times 10^{-10} \quad b_1 = 31.85 \times 10^{-10}$

$a_2 = 1.235 \times 10^{-17} \quad b_2 = -4.015 \times 10^{-17}$

$a_3 = -2.46 \times 10^{-26}$

This model is appealing in that it is evolved from accepted physics but that it yields Hubble's otherwise disturbing observational data to a fair approximation. The red-shifts are the classic Doppler effects. The non-homogeneity of the model is what should be expected from the known instability.

of homogeneous models. The theory appears to be sound and the fit of the model could be improved by further computation. There is considerable doubt, however, as to the wisdom of making a more detailed fit until the many questions as to the actual validity of Hubble's data are answered.

Having considered the primary data of red-shifts and counts, let us now turn to the secondary data of density and age. The density that must be assumed to fit the other data, namely $\rho_o = 3.3 \times 10^{-28}$ gms./c.c., is a reasonable one. Hubble (Ref. 16, p. 76) has estimated the order of magnitude of the averaged-out density of the luminous matter within the universe to be 10^{-30} gms./c.c. Sinclair Smith's (Ref. 25) measurements upon the Virgo cluster would indicate that the order of magnitude for the averaged-out density of all matter within the universe must be about 2×10^{-28} gms./c.c. The discrepancy between these two estimations must be due to a large part of the total mass of the universe being non-luminous. Perhaps it is scattered as meteoric material between the nebulae. Zwicky's (Ref. 26) measurements upon the Coma cluster would indicate an even higher value of about 4×10^{-28} gms./c.c. for the magnitude of the averaged-out density. Thus the value which must be assumed to make

our model fit is an entirely reasonable one. On the other hand, a more detailed study of the dynamics of the nebular clusters by the astronomers would be quite welcome since the density is a rather sensitive factor in our model. A slightly different density would produce considerably different behavior.

The situation for the calculated age of the universe is not so pleasant. It is about two-thirds of the most reasonable value. As indicated in the discussion at the beginning of this chapter, there are at least two reasons why the calculated age of the universe could be expected to be in error. First, the assumption of a central position for the observer is manifestly false. A non-central position for the observer would alter both R and T by considerable amounts. Second, all pressure effects have been neglected. Our model is in an early stage of evolution as shown by the small present value of the parameter $\Theta_o = 1.3603$. We would expect pressure effects to have been appreciable for a large fraction of its history. If these neglected pressure terms included a viscous-like force, we would expect our calculated age to be too small.

The age of the universe is not a well established datum as only a rough order of magnitude is given by the presently known data. The most common assumption is that

it lies between 2×10^9 years and 3×10^9 years. The lower limit is set by the radioactive determinations upon the oldest rock of the Earth. This lower limit is about 2×10^9 years. The same figure is found for the upper limit of radioactive determinations upon meteorities. This would appear to be significant since it has been estimated that 60% of the meteorities originate within interstellar space. The upper limit of roughly 10^{10} years is set by the calculated relaxation times for stellar and nebular clusters. Further rough supporting evidence is found in the continued existence of certain types of super-luminous stars and in the observed failure of equipartition of energy within our own galaxy. The most reasonable present value for the age of the universe would appear to be about 2×10^9 years. This assumes that the Earth was produced at the time of the singularity and formed its crust by sweeping up the meteoric material which was present with a high density at that epoch. This is in agreement with several present theories as to planetary evolution and would account for the known small eccentricity of the Earth's present orbit.

In concluding this chapter it must once more be emphasized that the material presented here is illustrative rather than definitive. We wish to show the

practical value of the seemingly abstruse theory of the first four chapters. The formulae developed for the red-shift, counts, and magnitudes, as well as the procedures utilized in roughly fitting this tenative model to the observational data, are the important elements. The precise figures of the various values evolved are not too consequential. They can be bettered by obvious, albeit Sisyphean labor.

References

1. R. C. Tolman, Relativity, Thermodynamics, and Cosmology, (Oxford, 1934)

2. R. C. Tolman, "Effect of Inhomogeneity on Cosmological Models", Proc. Nat. Acad. Sc., Vol. 20, pp. 169-176 (1934)

3. G. Lemaitre, Ann. de la Soc. Scient. de Bruxelles, Vol. A53, p. 51 (1933)

4. H. P. Robertson, "Relativistic Cosmology", Rev. of Mod. Phys., Vol. 5, pp. 62-90 (1933)

5. E. T. Copson, An Introduction to the Theory of Functions of a Complex Variable, (Oxford, 1935)

6. E. Jahnke & F. Emde, Funktionentafeln mit Formeln und Kurven, 3rd rev. ed., (Teubner, 1938)

7. E. Madelung, Die Mathematischen Hilfsmittel des Physikers, (Springer, 1936)

8. W. de Sitter, "The Expanding Universe. Discussion of Lemaitre's Solution of the Equations of the Inertial Field", Bull. Astron. Inst. Neth., Vol. 5, pp. 211-218 (1930)

   W. de Sitter, "Further Remarks on the Astronomical Consequences of the Theory of the Expanding Universe", Bull. Astron. Inst. Neth., Vol. 5, pp. 274-276 (1930)

   W. de Sitter, "Some Further Computations Regarding Non-static Universes", Bull. Astron. Inst. Neth., Vol. 6, pp. 141-145 (1931)

9. A. Einstein, The Meaning of Relativity, (Princeton Univ. Press, 2nd ed., (1945)

10. A. Einstein & E. G. Straus, "The Influence of the Expansion of Space on the Gravitation Fields Surrounding the Individual Stars", Rev. of Mod. Phys., Vol. 17, pp. 120-124 (1945)

11. P. G. Bergmann, Introduction to the Theory of Relativity, (Prentice-Hall, 1946)

12. C. Gregory, "On a Supplement to the Field Equations with an Application to Cosmology", Phys. Rev., Vol. 67, pp. 179-184 (1945)

13. G. Lemaitre, "Note on de Sitter's Universe", J. Math. and Phys. (M.I.T.) Vol. 4, pp. 188-192 (1925)

14. H. P. Robertson, "On Relativistic Cosmology", Phil. Mag., Vol. 5, pp. 835-848 (1928)

15. E. Hubble, "Effects of Red Shifts on the Distribution of Nebulae", Astrophys. J., Vol. 84, pp. 517-554 (1936)

16. E. Hubble, "The Distribution of Extra-Galactic Nebulæ", Astrophys. J., Vol. 79, pp. 8-76 (1934)

17. E. Hubble, "The Luminosity Function of Resolved Nebulae as Indicated by Their Brightest Stars", Astrophys. J., Vol. 84, pp. 158-179 (1936)

18. E. Hubble, "The Luminosity Function as Indicated by Residuals in Velocity-Magnitude Relations", Astrophys. J., Vol. 84, pp. 270-295 (1936)

19. E. Hubble, "The Surface Brightness of Threshold Images", Astrophys. J., Vol. 76, pp. 106-116 (1932)

20. E. Hubble, "The Motion of the Galactic system among the the Nebulae", Jour. Frank. Inst., Vol. 228, p. 131 (1939)

21. M. Wyman, "Equations of State for Radially Symmetric Distributions of Matter", Phys. Rev., Vol. 70, pp. 396-400 (1936)

22. G. Gamow, "Expanding Universe and the Origin of Elements", Phys. Rev., Vol. 70, pp. 572-573 (1946)

23. E. Hubble & R. C. Tolman, "Two Methods of Investigating the Nature of the Nebular Red-Shift", Astrophys. J., Vol. 82, pp. 302-337 (1935)

24. A. Fletcher, "The Number of Galaxies per Unit Volume", Mon. Not. R. A. S., Vol. 106, pp. 121-123 (1946)

25. S. Smith, "The Mass of the Virgo Cluster", Astrophys. J., Vol. 83, pp. 23-30 (1936)

26. F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln", Helv. Phys. Acta, Vol. 6, pp. 110-127 (1933)