BASIC CONCEPTS UNDERLYING SINGULAR PERTURBATION TECHNIQUES*

P. A. LAGERSTROM† AND R. G. CASTEN‡

Summary. In many singular perturbation problems multiple scales are used. For instance, one may use both the coordinate x and the coordinate $x^* = \varepsilon^{-1}x$. In a secular-type problem x and x^* are used simultaneously. This paper discusses layer-type problems in which x^* is used in a thin layer and x outside this layer. Assume one seeks approximations to a function $f(x, \varepsilon)$, uniformly valid to some order in ε for x in a closed interval D. In layer-type problems one uses (at least) two expansions (called inner and outer) neither of which is uniformly valid but whose domains of validity together cover the interval D. To define "domain of validity" one needs to consider intervals whose endpoints depend on ε. In the construction of the inner and outer expansions, constants and functions of ε occur which are determined by comparison of the two expansions "matching." The comparison is possible only in the domain of overlap of their regions of validity. Once overlap is established, matching is easily carried out. Heuristic ideas for determining domains of validity of approximations by a study of the corresponding equations are illustrated with the aid of model equations. It is shown that formally small terms in an equation may have large integrated effects. The study of this is of central importance for understanding layer-type problems. It is emphasized that considering the expansions as the result of applying limit processes can lead to serious errors and, in any case, hides the nature of the expansions.

1. Introduction. This paper is a survey of some essential ideas used in solving singular perturbation problems. It is not a survey of the literature on the subject; even for the problems discussed here no attempt has been made to make the references complete. It is not a survey of results or techniques—for this the reader is referred to Van Dyke (1964) and Cole (1968)—nor does it discuss exact mathematical results as does the book by Wasow (1965). Its main aim is to show the fundamental heuristic ideas which underlie certain techniques and which may be used in expanding these techniques.

The problems to be discussed are what in this paper will be called layer-type problems. Central concepts to be discussed are those of matching of asymptotic expansions, domains of validity of such expansions, overlap, formal validity of equations, limit processes, large integrated effects of formally small terms, switchback, etc. Many of these concepts were introduced by Saul Kaplun in the mid-fifties. Kaplun's analysis of concepts underlying older techniques and introduction of new concepts led to a much deeper understanding of older techniques and consequently to an extension of those techniques.

In § 2 a very simple physical example is used to show how singular perturbation problems arise, and the difference between secular-type and layer-type problems is illustrated. In § 3 various basic concepts are introduced. The motivation for

these concepts will be given in § 4, § 5 and § 6 which discuss in detail three model equations of increasing complexity. These equations were chosen to illustrate how various concepts and heuristic ideas are used in constructing asymptotic expansions.

Fluid dynamics has played an essential role in developing perturbation methods, similar to the role celestial mechanics played during the last century. Actually the first model equation (4.1) is similar to the one introduced by Friedrichs (1953) to illustrate Prandtl's boundary-layer theory for fluid flow at large Reynolds numbers, while the second and third model equations were introduced by one of the authors (see Lagerstrom (1961)) to illustrate Kaplun's analysis of fluid flow at small Reynolds numbers. However, at the same time the discussion in § 4, § 5 and § 6 aims at presenting the basic features of any typical layer-type problem.

Section 7 deals directly with problems from fluid dynamics. It tries to show how in many concrete physical problems the main body of perturbation ideas is applicable but in addition some special new ideas are needed. Fluid dynamics is still the source of many challenging perturbation problems.

There are many perturbation problems which properly may be called singular but which are not of the layer type. A large class of such problems are of the secular type (see § 2). For such problems the reader is referred to Kevorkian (1966) which discusses them in a systematic manner with many examples and to Cole (1968). The method of straining coordinates, as used by Whitham and Lighthill, is also omitted from the discussion; references are given in Van Dyke (1964).

2. A simple physical example. Before attempting any general description of what a singular perturbation problem is we shall start with a very elementary physical example which will show how both secular-type problems and layer-type problems may arise.

2.1. Time scales. We consider a linear spring-mass system with mass m, damping coefficient 2β and spring constant k. The equation is then

$$(2.1) \qquad m\frac{d^2 y_d}{(dt_d)^2} + 2\beta\frac{dy_d}{dt_d} + ky_d = 0,$$

where the amplitude y_d is measured in a dimensional length scale and t_d is the dimensional time. We know that the answer may be expressed in nondimensional terms. As nondimensional amplitude we may take

$$(2.2) \qquad y = y_d/L,$$

where L is a typical amplitude (of dimensional length). To find a nondimensional time scale we first observe that there are two independent ways of forming parameters of dimension time from the original parameters. We may define

$$(2.3) \qquad T_1 = \sqrt{m/k}$$

which is the inverse of the frequency. Another time parameter is

$$(2.4) \qquad T_2 = \beta/k.$$

From these we may form further time parameters

(2.5) $$T = T_1^a T_2^b, \quad a + b = 1.$$

One way of nondimensionalizing the time-coordinate is to define

(2.6) $$t = t_d/T_1.$$

Then (2.1) reads

(2.7a) $$\frac{d^2 y}{dt^2} + 2\varepsilon \frac{dy}{dt} + y = 0,$$

where

(2.7b) $$\varepsilon = \beta/\sqrt{mk} = T_2/T_1.$$

2.2. A secular-type problem. Suppose now we have a system of "small damping." Expressed in nondimensional form this means that

(2.8) $$\beta \ll \sqrt{mk} \quad \text{or, equivalently,} \quad \varepsilon \ll 1.$$

Making assumption (2.8) we study how one may solve (2.7) by perturbation techniques. We take as initial conditions

(2.9) $$y(0) = 0, \quad (dy/dt)_{t=0} = 1.$$

Because of (2.8) it seems natural to assume a first approximation y_0 satisfying

(2.10) $$d^2 y_0/dt^2 + y_0 = 0$$

and the initial conditions (2.9). Then

(2.11) $$y_0 = \sin t.$$

If we assume that y has an asymptotic expansion

(2.12) $$y \sim \sum_{k=0}^{\infty} \varepsilon^k y_k(t),$$

then y_1 must satisfy

(2.13a) $$d^2 y_1/dt^2 + y_1 = -2 \cos t$$

and

(2.13b) $$y_1(0) = y_1'(0) = 0.$$

Hence

(2.14) $$y_1 = -t \sin t$$

and

(2.15) $$y = y_0 + \varepsilon y_1 + O(\varepsilon^2) = \sin t - \varepsilon t \sin t + O(\varepsilon)^2.$$

We now ask whether (2.15) is a "good approximation." The answer obviously depends on the time interval considered. Consider the time interval

(2.16) $$D = [0, \tau].$$

If $\tau = O(1)$, then, as is easily checked, there exists a constant K such that

$$(2.17) \qquad\qquad |y - (y_0 + \varepsilon y_1)| \leqq K\varepsilon^2$$

for t in D. The error is then uniformly small and of the order formally indicated. However, if τ is large, say $\tau = O(1/\varepsilon)$, then (2.17) no longer holds. Using an obvious terminology we say that the formal order of $-\varepsilon t \sin t$ is ε but that its actual order for $t = O(1/\varepsilon)$ is unity. Thus $y_0 + \varepsilon y_1$ is uniformly valid to order ε^2 on the interval $[0, 1]$ but not on the interval $[0, 1/\varepsilon]$. This situation is the essence of a singular perturbation problem: The "straightforward" perturbation (2.12) fails to be uniformly valid.

A comparison with the exact solution shows that the approximation

$$(2.18) \qquad\qquad y \approx e^{-\varepsilon t} \sin t$$

is good to order unity on the interval of length $1/\varepsilon$. It is clear that the term (2.14) results from expansion of the exponential function in powers of εt. No partial sum of this expansion is valid to order ε if t is order ε^{-1}. Thus the assumption that a uniformly valid expansion could have the form (2.12) is wrong. A term such as (2.14) is called a secular term. Its occurrence is due to the resonant forcing function in (2.13a). However, in the present problem this resonance is not an actual physical phenomenon but a false mathematical resonance due to an improper perturbation method. We shall refer to problems of the type discussed here as secular-type problems. A systematic perturbation procedure for handling such problems must show how secular terms can be avoided. Such techniques are beyond the scope of the present paper. They are discussed extensively in Kevorkian (1966) and Cole (1968). We shall only draw attention to the following. In (2.18), $\sin t$ varies with $t = t_d/T_1$ whereas the exponential varies an order of magnitude more slowly, with $\varepsilon t = [(T_2/T_1)(t_d/T_1)] = t_d/T_3$, where

$$(2.19) \qquad\qquad T_3 = T_1^2 T_2^{-1} = m/\beta.$$

We are thus using two time scales of different orders of magnitude; the "normal" T_1 and the very long T_3. Correspondingly, we use two nondimensional time variables, t and εt. The first one is the "fast" variable and the second one the "slow" variable.

2.3. A layer-type problem. We shall now consider the other extreme of very large damping, or very small mass, i.e.,

$$(2.20) \qquad\qquad mk \ll \beta^2.$$

We may then write (2.1) as

$$(2.21) \qquad\qquad \eta\frac{d^2 y}{ds^2} + 2\frac{dy}{ds} + y = 0$$

with

$$(2.22) \qquad\qquad s = \frac{t_d k}{\beta} = \frac{t_d}{T_2}, \qquad \eta = \frac{km}{\beta^2} = \frac{1}{\varepsilon^2} = \frac{T_1^2}{T_2^2}.$$

The small parameter is now η. As initial conditions we take

(2.23a) $$y = A,$$

(2.23b) $$dy/ds = B \quad \text{at} \quad s = 0.$$

If we now assume an expansion of the form

(2.24) $$y \approx \sum_{k=0}^{\infty} \eta^k y_k(s),$$

we notice that there is something wrong even before we start solving the problem. The leading term $y_0(s)$ would obey the first order equation

(2.25) $$2\frac{dy_0}{ds} + y_0 = 0$$

and hence cannot be made to satisfy both initial conditions (2.23).

From the exact solution we find that

(2.26a) $$y \approx Ae^{-s/2} - \tfrac{1}{4}(A + 2B)\,e^{-2s/\eta} \cdot \eta,$$

(2.26b) $$dy/ds \approx -\tfrac{1}{2}Ae^{-s/2} + \tfrac{1}{2}(A + 2B)\,e^{-2s/\eta}$$

is an approximation to order unity in any interval $[0, s]$. We note that neglecting the term of formal order η in y one obtains $Ae^{-s/2}$ which satisfies (2.25) and the initial condition (2.23a). It gives, however, the wrong velocity at $s = 0$ and hence also for small values of s. The second term is needed to correct this mistake.

We see again that a correct approximation involves the use of two time scales and two nondimensional time variables:

(2.27a) $$s = t_d/T_2; \qquad T_2 = \beta/k,$$

(2.27b) $$s/\eta = t_d/T_3; \qquad T_3 = T_1^2 T_2^{-1} = m/\beta.$$

If we consider T_2 "normal," then T_3 is a very short scale. Both scales must be used in order to express a uniformly valid approximation.

The two perturbation problems considered above have the following in common: In both cases the assumption that the expansion in ε is a simple power series in ε leads to approximations which are not uniformly valid. In the first case the leading term of such a power series is valid to order unity in finite intervals. In the second it is valid everywhere except in a thin layer very near the origin where the mistake in the velocity is of order unity. We note that this correction term proportional to $e^{-2s/\eta}$ changes very rapidly in the sense that its changes are of order unity when the changes in the variable s are of order η. In each case one must use two time scales of different order of magnitude in order to obtain the solution. Thus in both problems we use a *method of multiple scales* (in the present case two time scales). In layer-type problems one scale is used in one region (which will later on be referred to as "outer") and the other scale, much smaller than the first, is used in another region which is a very thin layer (which will later on be referred to as "inner"). In secular-type problems both scales are used simultaneously. In our particular example there is a carrier wave, $\sin t$, whose

amplitude varied slowly as $e^{-\varepsilon t}$. Note that while the secular terms were due to incorrect perturbation methods there is a genuine secular effect; i.e., the damping is noticeable only if one studies the system over a very long time.[1] We notice that the existence of two independent time scales of different orders of magnitude gives one the small parameter ε and hence permits one to use perturbation methods. However, at the same time the existence of two such independent scales leads to the question of which scale should be used to nondimensionalize the time variables. Actually, we saw that in each case it was necessary to use both, i.e., the problems are singular. In general, as has been emphasized by Van Dyke (1964, p. 80), when the small parameter ε is a ratio of two lengths (or times) a perturbation scheme is likely to be singular. Of course, there are many singular perturbation problems which do not fit into this category. An example is flow at very high Mach number, which is discussed extensively in the literature (see, e.g., Van Dyke (1964)).

3. General concepts in perturbation problems. In this section we shall discuss some general concepts relating to asymptotic expansions. These concepts will be needed in the detailed discussion of specific examples in subsequent sections.

3.1. Coordinates and parameters; approximations. In perturbation problems we consider a function $f(x_1, \cdots, x_n; \varepsilon_1, \cdots, \varepsilon_m)$ depending on two sets of complex- or real-valued arguments, one set $(x_1, \cdots, x_n)$ called "variables" and another set $(\varepsilon_1, \cdots, \varepsilon_m)$ called parameters. The distinction between variables and parameters is not inherent in the nature of the function but depends on the context. The distinction may be made on physical grounds; or if the function is defined implicitly by differential equations and boundary conditions, then derivatives with respect to the variables but not derivatives with respect to the parameters occur in the equations. In a perturbation problem the distinction arises from the way the x_j and ε_j are treated.[2] For simplicity we consider one variable x and one parameter ε. In a perturbation problem we seek approximations to $y(x, \varepsilon)$ which should be uniformly valid for all x in a closed domain D when ε is very close to some distinguished value which without loss of generality we may assume to be zero. We say that $f(x, \varepsilon)$ is an approximation to $y(x, \varepsilon)$ uniformly valid to order $\zeta(\varepsilon)$ if

$$(3.1) \qquad \lim_{\varepsilon \to 0} \frac{y(x, \varepsilon) - f(x, \varepsilon)}{\zeta(\varepsilon)} = 0 \quad \text{uniformly for } x \text{ in } D.$$

In this definition there is of course no distinction between which function is the approximation and which is "the real thing." In a perturbation problem the function $y(x, \varepsilon)$ is usually defined implicitly by various equations (differential, integral, functional, boundary conditions). The purpose of perturbation methods is to try to construct approximating functions $f(x, \varepsilon)$, normally by the study of simplified equations. Often one constructs a sequence of approximations as will be discussed below.

In a physical problem ε may be a parameter such that only nonnegative real values of ε are physically meaningful. It is known that for a thorough mathematical

---

[1] The expression "secular" originated in classical celestial mechanics, where a century was considered a long time unit as opposed to short time units such as year or month.

[2] See, however, the discussion of coordinate-type expansions in § 7.

study of $y(x, \varepsilon)$ it is desirable to let ε assume complex values. We shall repeatedly deal with cases for which $y(x, \varepsilon)$ has an essential singularity at $\varepsilon = 0$ in the complex ε-plane. In this case the limit in (3.1) may depend on the path along which ε approaches zero. Unless otherwise stated we shall assume below that ε *approaches zero through positive values*. The results will usually be valid in some wedge of the complex ε-plane surrounding these values. We shall, however, not discuss complex values explicitly.

3.2. Sequence of approximations. In solving a perturbation problem one usually constructs what is, in principle, the beginning of an asymptotic expansion of $y(x, \varepsilon)$ at $\varepsilon = 0$.

Let $\zeta_j(\varepsilon), j = 0, 1, \cdots$, be an asymptotic sequence (for $\varepsilon \downarrow 0$), i.e.,[3]

$$(3.2) \qquad \lim_{\varepsilon \downarrow 0} \zeta_{j+1}(\varepsilon)/\zeta_j(\varepsilon) = 0.$$

A sequence of uniformly valid asymptotic approximations to $y(x, \varepsilon)$ with respect to the sequence ζ_j is a sequence of functions $a_0(x, \varepsilon), a_1(x, \varepsilon), \cdots$ such that for any j,

$$(3.3) \qquad \lim_{\varepsilon \downarrow 0} \frac{y(x, \varepsilon) - a_j(x, \varepsilon)}{\zeta_j(\varepsilon)} = 0$$

uniformly in D. Thus each $a_j(x, \varepsilon)$ is an approximation uniformly valid to order ζ_j. We refer to the $\zeta_j(\varepsilon)$ as the *gauge functions* for this approximation scheme.

A very common method for constructing the $a_j(x, \varepsilon)$ is to find a sequence $b_j(x, \varepsilon)$ such that

$$(3.4) \qquad a_j(x, \varepsilon) = \sum_{k=0}^{j} b_k(x, \varepsilon).$$

We say that if (3.3) and (3.4) hold, the formal sum $\sum_{k=0}^{\infty} b_k(x, \varepsilon)$ is an asymptotic expansion of $y(x, \varepsilon)$ relative to the gauge functions $\zeta_j(\varepsilon)$, uniformly valid in D, and write

$$(3.5) \qquad y(x, \varepsilon) \simeq \sum_{k=0}^{\infty} b_k(x, \varepsilon).$$

For the basic properties of asymptotic expansions the reader is referred to standard literature, for instance, Wasow (1965) or Erdélyi (1956).

To summarize, a perturbation method as understood here is a method of constructing an asymptotic expansion, or the beginning of an asymptotic expansion of a function $y(x, \varepsilon)$. The partial sums of (3.5) will be approximations to $y(x, \varepsilon)$. Note, however, that the error term is an order estimate. We may rewrite (3.3) as

$$(3.6) \qquad y(x, \varepsilon) - a_j(x, \varepsilon) = o(\zeta_j(\varepsilon)).$$

To convert this into a numerical error estimate one needs to find functions $\zeta_j^*(\varepsilon)$, where $\zeta_j^* = o(\zeta_j)$, such that $|y(x, \varepsilon) - a_j(x, \varepsilon)| \leq \zeta_j^*(\varepsilon)$ for all x in D. The numerical aspect of the approximation scheme will not be dealt with here.

---

[3] To express relations such as (3.2) we shall often use the simpler notation $\zeta_{j+1} \ll \zeta_j$ or $\zeta_{j+1} = o(\zeta_j)$.

3.3. Regular and singular perturbations. It would be tempting to distinguish between regular and singular perturbation problems according to whether $y(x, \varepsilon)$ is regular or singular around $\varepsilon = 0$. This, however, leads to various complications. One of these is the following: If we introduce x^* and ε^* by[4]

(3.7a) $\qquad\qquad x^* = $ function of x and ε,

(3.7b) $\qquad\qquad \varepsilon^* = $ function of ε such that $\varepsilon^* = 0$ for $\varepsilon = 0$,

(3.7c) $\qquad\qquad f^*(x^*, \varepsilon^*) = f(x, \varepsilon)$,

then x^* is still a variable and ε^* still a parameter in the sense discussed above. It may happen that f is singular at $\varepsilon = 0$ but f^* is regular at $\varepsilon^* = 0$. A trivial example is, with

(3.8a) $\qquad\qquad D = \{x | x \geqq 0\}$,

(3.8b) $\qquad\qquad f(x, \varepsilon) = e^{-x/\varepsilon} + x$,

(3.8c) $\qquad\qquad x^* = x/\varepsilon, \qquad \varepsilon^* = \varepsilon$,

(3.8d) $\qquad\qquad f^*(x^*, \varepsilon^*) = e^{-x^*} + \varepsilon x^*$.

We say that a uniformly valid asymptotic expansion (3.5) is regular, with respect to the domain D and the gauge functions $\zeta_j(\varepsilon)$, if each $b_j(x, \varepsilon)$ may be written

(3.9) $\qquad\qquad\qquad b_j = \beta_j(\varepsilon) b_j(x)$,

that is, if (3.5) has the form

(3.10) $\qquad\qquad\qquad y(x, \varepsilon) \simeq \sum_{j=0}^{\infty} \beta_j(\varepsilon) b_j(x)$,

where the $\beta_j(\varepsilon)$ form an asymptotic sequence which may or may not coincide with the sequence of gauge functions $\zeta_j(\varepsilon)$. We do *not* require that the $\beta_j(\varepsilon)$ be powers of ε but admit, for instance,

$$\beta_0 = 1, \quad \beta_1 = \varepsilon \ln \varepsilon, \quad \beta_2 = \varepsilon, \quad \text{etc.}$$

Thus it is not necessary for $f(x, \varepsilon)$ to be regular at $\varepsilon = 0$. The two terms of the right-hand side of (3.8b) do obviously not constitute a regular expansion. Less obviously so, if $\zeta_j = \beta_j = \varepsilon^j$, then (3.8d) is not a regular expansion either since the term εx^* is not uniformly small in the domain D considered. Again we see that regularity in ε at $\varepsilon = 0$ is not necessarily a criterion for the possibility of a regular expansion.

A regular perturbation method is a systematic way of successively constructing the $\beta_j(\varepsilon)$ and $b_j(x)$, possibly preceded by a change of coordinates of the type (3.7).

A singular perturbation method may be defined in general as a method which is not regular. To make a positive statement, in layer-type problems we shall normally construct uniformly valid expansions by constructing expansions of the form (3.10) which however are not uniformly valid and which must be

---

[4] We shall denote the transformation (3.7) by the vague term "coordinate change."

complemented by one or several expansions

$$(3.11) \qquad y(x, \varepsilon) \simeq \sum \beta_j(\varepsilon) a_j(x^*) ,$$

where $x^* = $ function of ε and x is a new variable of different order of magnitude than x, say $x^* = x/\varepsilon$. The form is the same as that of (3.10) except that we use x^* as coordinate. Actually, as we shall see in § 6, one or several of the expansions may not be of the special form (3.10), (3.11) but will still be of the general form (3.5). Obviously, in some sense the various expansions together must furnish approximations uniformly valid in D. This important concept will be discussed below.

Strictly speaking one should not speak of singular perturbation *problems* but of singular perturbation *methods* or *techniques*. As an analogy consider a differential equation with boundary conditions of such a nature that Fourier analysis is a very natural and simple method for obtaining the solution. Still, we do not call this problem a Fourier problem. Similarly, there are perturbation problems for which it is natural to construct a singular asymptotic expansion. However, it might have been possible by a clever a priori coordinate change, to construct a regular asymptotic expansion. Or, as will be shown in § 7, a regular expansion may be constructed on the basis of the singular expansion. However, in this paper we shall continue to use the convenient term "singular perturbation problem."

3.4. Validity of an approximation. The aim of perturbation methods is to obtain approximations which are uniformly valid in the given domain D. It is typical of singular perturbation techniques (of the layer type) that one works with approximations which are valid only in more restricted domains. It will be necessary to define the concept of uniform domain of validity for such approximations. Thus we shall generalize (3.1).

As an example consider the functions

$$(3.12a) \qquad y(x, \varepsilon) = e^{-x/\varepsilon} + x + \varepsilon,$$

$$(3.12b) \qquad f(x, \varepsilon) = x + \varepsilon,$$

$$(3.12c) \qquad g(x, \varepsilon) = e^{-x/\varepsilon}.$$

We shall consider $y(x, \varepsilon)$ as "the exact solution" and treat $f(x, \varepsilon)$ and $g(x, \varepsilon)$ as approximations. The domain of the variable x will be the closed interval $[0, 1]$. We see that f is an approximation to y, uniformly valid to order unity on any interval $[x_0, 1]$ provided $0 < x_0 \leqq 1$, while it is not valid on $[0, 1]$ (and not uniformly valid on $[0, 1]$ even if one redefines $f(0, \varepsilon)$ to be $= 1$).

The statements just made seem to exhaust the description of where $f(x, \varepsilon)$ is a uniformly valid approximation. However, for the theory of matching it is essential that the notion of "uniform validity" be extended. Consider the continuum of intervals $[\sqrt{\varepsilon}, 1]$ where ε ranges in some open interval $0 < \varepsilon < \varepsilon_0 \leqq 1$. The maximum error in any such interval occurs at the left endpoint and is $e^{-1/\sqrt{\varepsilon}}$. Then it is clear that given any $\delta > 0$ one can find an ε_δ such that $|y - f| < \delta$ for all ε, $0 < \varepsilon < \varepsilon_\delta$ and all x in $[\sqrt{\varepsilon}, 1]$, i.e., in the vertically striped region of the (x, ε)-plane as shown in Fig. 3.1. The corresponding domain for an interval with
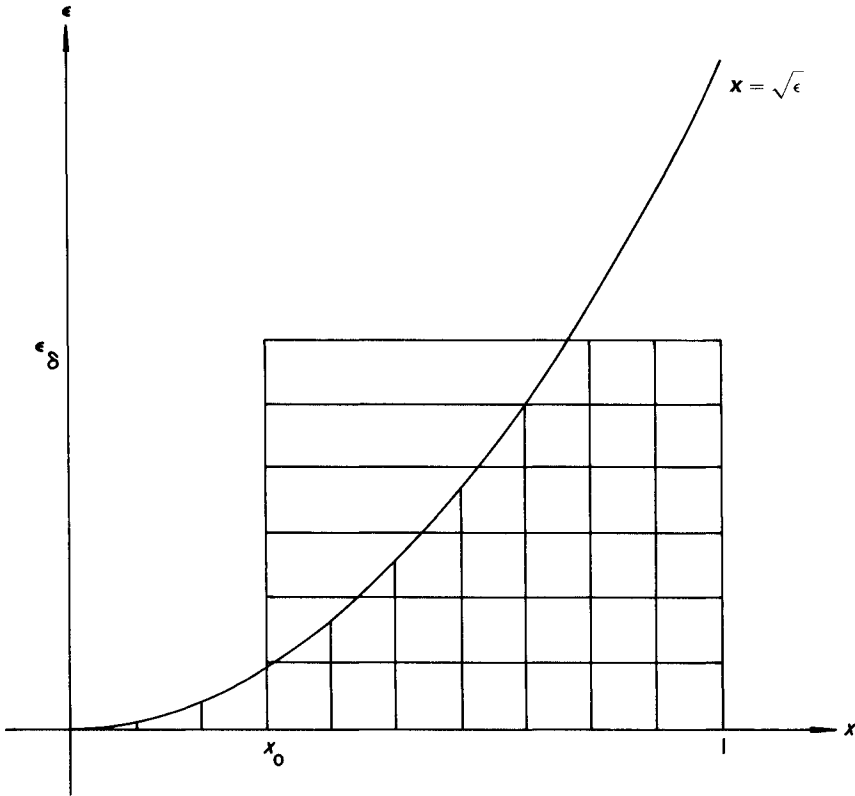
FIG. 3.1. *Domains of validity*

fixed endpoints would look like the horizontally striped rectangle in the figure. We shall simply refer to $[\sqrt{\varepsilon}, 1]$ as an interval, with a moving endpoint (rather than a continuum of intervals). It is natural to say that f is an approximation to y uniformly valid on the interval $[\sqrt{\varepsilon}, 1]$. We note that in this statement $\sqrt{\varepsilon}$ may be replaced by any function η such that $0 < \eta(\varepsilon) < 1$ for ε in some interval $(0, \varepsilon_0)$ provided that $\varepsilon \ll \eta$, i.e., $\lim (\varepsilon/\eta) = 0$ as $\varepsilon \downarrow 0$. Similarly we see that $g(x, \varepsilon)$ is an approximation to $y(x, \varepsilon)$ uniformly valid on any interval $[0, \eta]$ provided $\eta(\varepsilon) \ll 1$, and $0 < \eta(\varepsilon) < 1$ for ε sufficiently small.

The above examples suggest the following definitions.

(3.13) DEFINITION. A function $w(x, \varepsilon)$ *converges uniformly to zero with respect to $\mathscr{D}$* (a class of functions $\eta(\varepsilon)$) if for any two $v(\varepsilon)$ and $\mu(\varepsilon)$ in $\mathscr{D}$ it converges to zero uniformly on the interval $[\mu, v]$, i.e., for each $\delta > 0$ there exists an ε_δ such that $|w(x, \varepsilon)| < \delta$ for $\varepsilon < \varepsilon_\delta$ and $\mu(\varepsilon) \leqq x \leqq v(\varepsilon)$.

This definition needs considerable polishing although in practice no difficulties occur: For values of ε such that $v(\varepsilon) < \mu(\varepsilon)$, the interval $[\mu, v]$ is empty. The most convenient handling of this is to say that the error is zero on an empty interval (see Kaplun (1967)). Also, the interval $[\mu, v]$ may be partly outside the domain of definition of $w(x, \varepsilon)$. This can be easily handled by trimming the functions in a proper way. A more essential point is the following. In layer-type problems

there will be exceptional points where the machinery we are developing here will be needed. In the study of each such point we can always introduce a coordinate such that the point is at the origin. Furthermore, for such points the behavior of the function on one side may be radically different from its behavior on the other side (a classical example is $e^{-x/\varepsilon}$). In studying the behavior on one side we may use a coordinate which is positive on that side. For this reason we shall assume that functions $\eta(\varepsilon)$ are nonnegative. More precisely we shall assume that either $\eta(\varepsilon) > 0$ or that $\eta(\varepsilon) = 0$ over a whole interval $(0, \varepsilon_0)$. These conventions are not motivated yet but their purpose will be illustrated by the various examples given later. With these restrictions on our function $\eta(\varepsilon)$ we make the following definition.

(3.14) DEFINITION. A function $f(x, \varepsilon)$ is an approximation to the function $y(x, \varepsilon)$ uniformly valid to order $\zeta(\varepsilon)$ with respect to the function class $\mathscr{D}$ if $(y(x, \varepsilon) - f(x, \varepsilon))/\zeta(\varepsilon)$ converges to zero uniformly with respect to $\mathscr{D}$.

Thus in (3.13) we generalize the concept of uniform convergence on an interval to that of uniform convergence on a function class. Similarly (3.14) generalizes the concept of uniform validity on an interval to that of uniform validity on a function class.[5]

3.5. Domains of validity.

Returning to the example given by (3.12) we see that the domain of uniform validity (to order unity) of $f(x, \varepsilon)$ as an approximation to $y(x, \varepsilon)$ is the function class

$$(3.15a) \qquad \mathscr{D}_f = \{\eta | \varepsilon \ll \eta\}.$$

Similarly the domain of validity for the approximation $g(x, \varepsilon)$ is

$$(3.15b) \qquad \mathscr{D}_g = \{\eta | \eta \ll 1\}.$$

Here the functions $\eta(\varepsilon)$ are restricted as discussed above.

The following lemma is trivial.

(3.16) LEMMA. *Let $\mathscr{D}$ be a domain of validity of an approximation $f(x, \varepsilon)$ and let $\eta_1 < \eta_3 < \eta_2$ with η_1 and η_2 in $\mathscr{D}$. Then the domain of validity can be extended to include η_3. The same conclusion follows if $\eta_1 \ll \eta_3 \ll \eta_2$.*

Thus a domain of validity can always be extended to include functions in-between.

In most practical cases there will be a maximal domain of validity. When we talk about *the* domain of validity we shall mean the maximal one.

3.6. Order classes.

We note in our example that whether a function $\eta(\varepsilon)$ belongs to one of the domains (3.15a) or (3.15b) depends only on its order of magnitude. It is convenient to formalize this by the following definition.

(3.17a) DEFINITION. *The functions η and ζ are o-equivalent if $\lim \eta/\zeta$ as $\varepsilon \downarrow 0$ exists and is different from zero and infinity.*

---

[5] The definition of asymptotic expansion introduced earlier is a natural but trivial generalization of Poincaré's definition. On the other hand, the introduction of the concepts used in (3.13) and (3.14), due to Kaplun, is definitely nontrivial; these concepts will be basic tools in the study of the nature of matching.

This is obviously an equivalence relation. Hence we may define the equivalence class or order class (*o*-class)

(3.17b) $$\text{ord } \eta = \{\zeta | \zeta \ o\text{-equivalent to } \eta\}.$$

We may also define the following more inclusive relation.

(3.18a) DEFINITION. The functions η and ζ are *O-equivalent* if there exist constants A and B, $0 < A < B$, such that $A < \zeta/\eta < B$ for ε sufficiently small.

Similarly we define the *O*-class

(3.18b) $$\text{Ord } \eta = \{\zeta | \zeta \ O\text{-equivalent to } \eta\}.$$

There is a natural way of introducing a partial ordering among the equivalence classes:

(3.19a) $$\text{ord } \eta < \text{ord } \zeta \quad \text{and} \quad \text{Ord } \eta < \text{Ord } \zeta$$

both mean

(3.19b) $$\lim \eta/\zeta = 0, \quad \text{i.e., } \eta \ll \zeta.$$

The ordering is only partial: there may be equivalence classes which are incompatible, i.e., distinct but neither being greater than the other. Since $\text{ord } \eta_1 < \text{ord } \eta_2$ implies $\text{Ord } \eta_1 < \text{Ord } \eta_2$, the class $\text{Ord } \eta$ consists of pairwise incompatible *o*-classes.

The concept of order classes would probably play an important role in any rigorous theory of singular perturbations. Also, the structure of sets of such classes leads to various interesting problems. For further details the reader is referred to Kaplun (1967).

3.7. Overlap. The domains (3.15a) and (3.15b) are each half infinite intervals which in an obvious sense are open. The domains of validity have the intersection

(3.20a) $$\mathcal{D}_f \cap \mathcal{D}_g = \{\eta | \varepsilon \ll \eta \ll 1\}.$$

This is called the *overlap domain* of the two approximations (3.12b) and (3.12c). More precisely it is the overlap domain to *order unity*.

At the same time, together, they *cover* the entire interval under consideration. $\mathcal{D}_g$ includes the function $\eta \equiv 0$ and all "infinitesimal" values of x, i.e., all $x = x_\eta \cdot \eta$, $\eta = o(1)$. $\mathcal{D}_f$ includes all constant functions $\eta = a$, $0 < a \leqq 1$, and enough infinitesimals to overlap with $\mathcal{D}_g$. We then say that the union of the two domains is the complete interval

(3.20b) $$\mathcal{D}_f \cup \mathcal{D}_g = [0, 1].$$

We shall use the expression that the functions f and g of (3.12b), (3.12c) *together* give an approximation to the function y of (3.12a) which is uniformly valid to order unity on the interval $[0, 1]$.

3.8. Limit processes. An important tool in singular perturbation techniques is that of a limit process. We define the η-limit of a function $y(x, \varepsilon)$ as follows. First we introduce the variable x_η by

(3.21) $$\eta x_\eta = x.$$

Then the η-limit of f is

(3.22) $$\lim_\eta f = \lim f \quad \text{as } \varepsilon \downarrow 0 \quad \text{with } x_\eta \text{ fixed}.$$

Thus given the proper continuity of f we may rewrite f as a function of x_η and ε and then put $\varepsilon = 0$. To illustrate this concept consider $y(x, \varepsilon)$ as defined by (3.12a). Then

(3.23a) $$\lim_\eta y = e^{-x/\varepsilon} \quad \text{for } \eta = \varepsilon,$$

(3.23b) $$\lim_\eta y = 0 \quad \text{for } \eta = \sqrt{\varepsilon},$$

(3.23c) $$\lim_\eta y = x \quad \text{for } \eta = 1.$$

3.9. A fundamental lemma. The following lemma is trivial and its proof will be omitted. It is, however, fundamental for the theory of matching (cf. (3.16)).

(3.24) LEMMA. *Let $f(x, \varepsilon)$ and $g(x, \varepsilon)$ be approximations to $y(x, \varepsilon)$ and let $\mathscr{D}$ be an overlap domain to order $\zeta(\varepsilon)$. Let η_1 and η_2 be in $\mathscr{D}$ and assume $\eta_1 \ll \eta \ll \eta_2$. Then*

$$\lim_\eta \frac{f(x, \varepsilon) - g(x, \varepsilon)}{\zeta(\varepsilon)} = 0.$$

3.10. Matching. Let f and g be two asymptotic approximations to y. They are said to *match* if their domains of validity overlap. If for instance f and g have been found from approximate equations but g still involves one as yet undetermined constant, then one may determine this constant by *matching*, i.e., by comparing
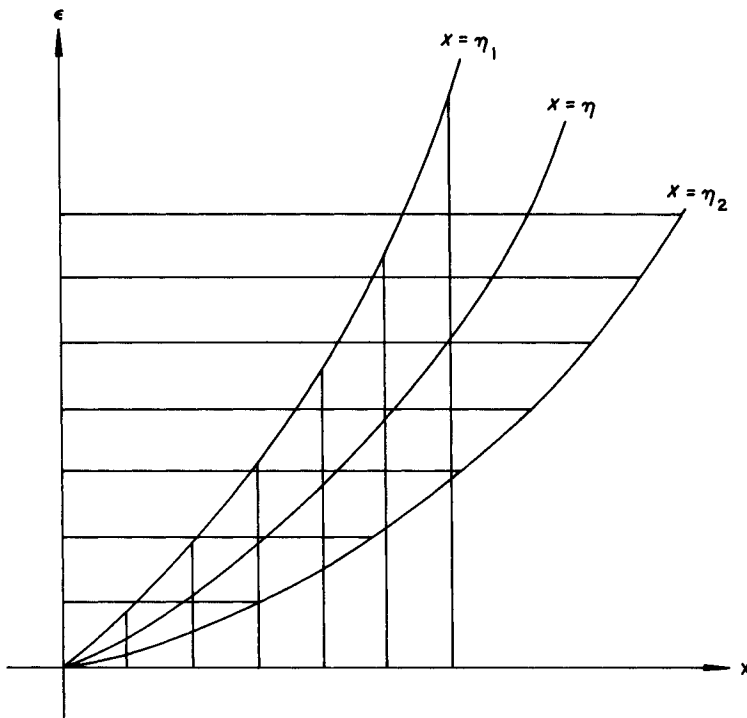

FIG. 3.2. *Matching in overlap domain*

them in their domain of overlap. In *principle*,[6] the matching may be carried out by applying (3.24). Taking the $\lim_n$ means letting ε approach zero on the curve $x = \eta(\varepsilon)$ as shown in Fig. 3.2. Here vertical stripes indicate part of the domain of validity of f and horizontal stripes refer to the domain of validity of g.

3.11. The extension theorem. Clearly, in order to use matching, we need criteria for determining as large a portion as possible of the domain of validity of an approximation. The following theorem, due to Kaplun, will play an essential role in determining domains of validity for the examples in § 5 and § 6.

(3.25) THEOREM. *Let $w(x, \varepsilon)$ tend to zero uniformly with respect to a class of functions $\mathcal{D}$ which is maximal and let all functions of an o-class Y be in $\mathcal{D}$. Then there exists an o-class X in $\mathcal{D}$ such that $X < Y$.*

Proof. Assume first that $Y = \text{ord } 1$. To fix the ideas, consider the interval $0 \leqq x \leqq 1$. By assumption, given a constant $c > 0$ and a $\delta > 0$ one may find an $\varepsilon_\delta > 0$ such that $w(x, \varepsilon) < \delta$ for $c \leqq x \leqq 1, 0 < \varepsilon \leqq \varepsilon_\delta$. The trick of the proof is to use this property for a sequence of constants c_n and a sequence δ_n, both tending to zero. One may, for instance, put $\delta_n = c_n = 1/n$. Since all constants $1/n$ belong to ord 1, we may, for each n, pick an ε_n such that $w < 1/n$ for $1/n \leqq x \leqq 1$; $0 < \varepsilon \leqq \varepsilon_n$. We now define a function $\eta(\varepsilon)$ by the requirement $\eta(\varepsilon_{n+1}) = 1/n$ and assume that η is a linear function of ε between any two adjacent values ε_k and ε_{k+1}. We can now find an X such that $\text{ord } \eta < X < Y$. It is easily seen that any such X is in $\mathcal{D}$. In the general case, when $Y = \text{Ord } \zeta$, we replace x by x/ζ in the above argument. Obviously the theorem can be proved with " $<$ " replaced by " $>$ ". Also, the proof shows that the whole O-class corresponding to X is in $\mathcal{D}$.

The extension theorem shows that a maximal domain of validity cannot have an infimum or a supremum. It is open in the sense that for each O-class Y there exist O-classes X and Z in the domain such that $X < Y < Z$.

We shall now illustrate the technique of singular perturbations and the underlying concepts and heuristic ideas by several model examples.

4. First model equation.
4.1. Formulation of problem. As Example I we take a function $y(x, \varepsilon)$ defined implicitly by the differential equation

$$(4.1a) \qquad \varepsilon \frac{d^2 y}{dx^2} + \frac{dy}{dx} - a - 2bx = 0$$

and the boundary conditions

$$(4.1b) \qquad y(0, \varepsilon) = 0,$$

$$(4.1c) \qquad y(1, \varepsilon) = 1.$$

Here a and b are constants, independent of ε.

By solving the above equation we find the explicit expression

$$(4.2) \quad y = (1 - a - b + 2\varepsilon b)\frac{1 - e^{-x/\varepsilon}}{1 - e^{-1/\varepsilon}} + bx^2 + ax - 2\varepsilon bx = \tilde{y} + O(e^{-1/\varepsilon}),$$

---

[6] How it is done in practice will be discussed later.

where $\tilde{y}$ is obtained from y by omitting the term $e^{-1/\varepsilon}$ in the denominator of the first term of y.

We shall now show how one may construct an asymptotic expansion of y by a method of singular perturbations applied to (4.1a), (4.1b), (4.1c). The results may be analyzed by comparison with (4.2). We shall seek an asymptotic expansion uniformly valid on the interval $[0, 1]$ as ε tends to zero through positive real values. As gauge functions we shall use the sequence (ε^k), $k = 0, 1, \cdots$. A function $\zeta(\varepsilon)$ is called transcendentally small if it is of smaller order of magnitude than all gauge functions. The function $e^{-1/\varepsilon}$ is then transcendentally small relative to our gauge functions. Using only the ε^k as gauge functions we cannot expect our construction to distinguish between y and $\tilde{y}$ in (4.2).

4.2. Perturbation methods. In a regular perturbation method we would assume that $y(x, \varepsilon)$ has an asymptotic expansion

$$(4.3) \qquad y \simeq \sum_{k=0}^{\infty} \varepsilon^k f_k(x).$$

(The ε^k might of course have to be replaced by less simple functions of ε.) By insertion of (4.3) into the equation and boundary conditions for y and grouping terms of the same order in ε one finds an equation and boundary conditions for f_0, then an equation and boundary conditions for f_1 which may assume that f_0 is known, etc. In our case the equation for f_0 is

$$(4.4a) \qquad df_0/dx - a - 2bx = 0.$$

This equation is of first order. Thus it will be impossible to satisfy both boundary conditions (4.1b) and (4.1c). For reasons to be discussed later (after (4.52)) we impose the second boundary condition

$$(4.4b) \qquad f_0(1) = 1.$$

This determines f_0 to be

$$(4.5) \qquad f_0 = 1 - a - b + ax + bx^2.$$

The value of f_0 at $x = 0$ is $1 - a - b$. Excluding the accidental case of $a + b = 1$ we find that $f_0(0) \neq 0$. Thus while f_0 is, in some sense to be determined, an approximation to the exact solution y, the error at $x = 0$ is of order unity. Hence f_0 is *not uniformly valid* near $x = 0$. We thus have a singular perturbation problem. To get an approximation which is uniformly valid we shall have to complement f_0 by an approximation valid near $x = 0$. We shall now show how such an approximation is constructed. The justification of the method will follow later. We introduce the stretched variable x^* defined by

$$(4.6) \qquad \varepsilon x^* = x.$$

Expressed in terms of this variable, (4.1a) becomes

$$(4.7) \qquad \frac{d^2 y}{dx^{*2}} + \frac{dy}{dx^*} - \varepsilon a - 2\varepsilon^2 bx^* = 0.$$

The introduction of this variable is based on the assumption that except very near to the origin $f_0(x)$ is a good approximation. More precisely, there is a layer of thickness ε when y changes from the value 0 at $x = 0$ to values close to $f_0(x)$. Because of this rapid change the term dy/dx does not have its "formal" order, which is unity, but is actually of order $1/\varepsilon$. The introduction of the variable formalizes this order determination since $dy/dx = (1/\varepsilon)dy/dx^*$.

We now assume that the approximation to $y(x, \varepsilon)$, valid near the origin, has an asymptotic expansion of the form

$$(4.8) \qquad y \simeq g_0(x^*) + \varepsilon g_1(x^*) + \cdots.$$

From (4.7) we see that g_0 must satisfy the equation

$$(4.9a) \qquad \frac{d^2 g_0}{dx^{*2}} + \frac{dg_0}{dx^*} = 0.$$

Since this approximation is supposed to be valid near the origin, the boundary condition at $x = 0$ must be satisfied, i.e.,

$$(4.9b) \qquad g_0(0) = 0.$$

In determining f_0 we could only impose *one* boundary condition on f_0 and had to relinquish the condition at $x = 0$. We are now faced with the problem of having too few boundary conditions: The second order equation (4.9a) requires two boundary conditions and while it is natural to impose (4.9b), it does not make sense to impose the condition at $x = 1$ since g_0 is supposed to be an approximation only for x near the origin.

4.3. Formal matching. We are now employing a technique of singular perturbations in that we are using two scales. The approximation $f_0(x)$ is obtained by using a scale of order unity (relative to the variable x). The approximation $g_0(x^*)$ is obtained by using a scale of order ε. This scale is used only near the value $x = 0$. We say that a *boundary layer* (or layer of rapid transition) of thickness ε exists at the origin. In this layer the solution $f_0(x)$ is not valid, but we use instead the *boundary-layer solution* $g_0(x^*)$. Since the scale ε is used only in this narrow region we are dealing with a layer-type problem. An essential feature of a layer-type singular perturbation technique is that the missing boundary condition for the boundary-layer solution is supplied by *matching* with the solution $f_0(x)$. In our case we may use as a formal matching condition

$$(4.9c) \qquad g_0(\infty) = f_0(0) \equiv 1 - a - b.$$

This condition may seem very surprising. For instance, g_0 is supposed to make sense only for x small and in (4.9c) we are using it at $x^* = \infty$ which would also mean $x = \infty$ if ε were fixed. At the same time we are comparing $g_0(\infty)$ with the value of $f_0(x)$ for $x = 0$. At present we only use the very vague justification that for ε very small there are values of x near zero such that x/ε is near infinity and where both f_0 and g_0 are valid approximations. One of the main objects of the present paper is to analyze and formalize this justification. This will lead to a general method of matching which, in the present special case, will imply (4.9c).

Using the boundary condition (4.9b) we find

(4.10a)
$$g_0 = C(1 - e^{-x^*}).$$

The matching condition (4.9c) then gives

(4.10b)
$$C = 1 - a - b.$$

4.4. Higher order approximations. We have seen that the expansion (4.3) cannot be expected to be valid at the origin. We call it an *outer* expansion, while the expansion (4.8), expected to be valid near the origin, is called the *inner* expansion.[7] The first terms of each expansion have been determined and we shall now find the higher order terms. From (4.3) we find that f_1 satisfies

(4.11a)
$$\frac{df_1}{dx} = -\frac{d^2f_0}{dx^2} = -2b.$$

Again we require that $f_0 + \varepsilon f_1$ satisfy the boundary condition at $x = 1$ which gives

(4.11b)
$$f_1(1) = 0,$$

and hence

(4.12)
$$f_1 = 2b - 2bx.$$

From (4.7) we find that g_1 satisfies the equation

(4.13a)
$$\frac{d^2g_1}{dx^{*2}} + \frac{dg_1}{dx^*} - a = 0.$$

The boundary condition at $x = 0$ gives

(4.13b)
$$g_1(0) = 0.$$

Hence,

(4.14)
$$g_1 = C_1(1 - e^{-x^*}) + ax^*,$$

where C_1 is a constant to be determined by matching. Obviously, the simple matching principle (4.9c) will not work. We replace it by the more general, but rather vague rule, which is a generalization of the justification given for (4.9c):

(4.15) $f_0 + \varepsilon f_1$ for x small must agree to order ε with $g_0 + \varepsilon g_1$ for x^* large.

---

[7] In this paper we apply the terms "outer" and "inner" to expansions, limits, variables and regions. The reason for the use of the terms "outer" and "inner" is historical. Its origin dates back to the classical boundary-layer problem of viscous flow past an object. (See § 7.3.) Actually the terminology will be incorrect for the flow inside a domain also discussed in § 7. (In this case the inner region surrounds the outer region.) We shall, however, follow the historical usage. Thus the inner region will always be small, in our case of order ε, relative to the outer region.

In the discussion after (4.31) we shall present an equivalent way of characterizing the outer limit as the first principal limit. In a sense to be described it takes precedence over the second principal limit which is the inner limit.

See also § 4.14.

For x^* large the exponential terms are negligible and

(4.16a) $$g_0 + \varepsilon g_1 \simeq 1 - (a + b) + \varepsilon(C_1 + ax^*).$$

For x small we have

(4.16b) $$f_0 + \varepsilon f_1 \simeq ax + 1 - (a + b) + 2b\varepsilon.$$

Here we have neglected terms of second order in x and ε, i.e., $bx^2 - 2\varepsilon bx$. Thus (4.16a) and (4.16b) agree if we put $C_1 = 2b$, i.e.,

(4.17) $$g_1 = 2b(1 - e^{-x^*}) + ax^*.$$

Continuing the same reasoning we find

(4.18) $$f_k = 0, \quad k > 1,$$

(4.19a) $$g_2 = bx^{*2} - 2bx^*,$$

(4.19b) $$g_k = 0, \quad k > 2.$$

4.5. Comparison with exact solution. By asymptotic construction we have obtained all the terms of the inner and the outer expansion. In what sense have we obtained an asymptotic approximation to the exact solution? To answer this we need the concept of domain of validity introduced in § 3.

First we consider f_0. Let x be in the interval $[\eta, 1]$, where η is some function of ε. Since

(a) $$|y - f_0| = |1 - a - b| \, e^{-x/\varepsilon} + \varepsilon \cdot \text{bounded function},$$

the maximal error occurs at the left endpoint where $x = \eta$. The error tends uniformly to zero in this interval provided $\varepsilon \ll \eta$. Thus,

(4.20a) the function f_0 is an approximation to y uniformly valid to order unity on the domain $\mathcal{D}_1 = \{\eta | \varepsilon \ll \eta\}$.

Similarly,

(4.20b) $f_0 + \varepsilon f_1$ is uniformly valid to all orders ε^k on $\mathcal{D}_1$.

It is clear that validity to order ε is the essential feature. Validity to higher orders would be spoiled by a simple modification, such as introducing a term cx^2 into the equation (4.1a) or by making $y(1, \varepsilon) = 1 + \varepsilon^2$.

The validity of the inner expansion is as follows:
Domain of validity of g_0 to order unity is

(4.21a) $$\mathcal{D}_2 = \{\eta | \eta \ll 1\}.$$

Domain of validity of $g_0 + \varepsilon g_1$ to order ε is

(4.21b) $$\mathcal{D}_3 = \{\eta | \eta \ll \sqrt{\varepsilon}\}.$$

The reason for the smaller domain in the second case is that

(b) $$|(y - g_0 - \varepsilon g_1)\varepsilon^{-1}| = bx^2/\varepsilon - 2bx.$$

Finally, $g_0 + \varepsilon g_1 + \varepsilon^2 g_2$ is the exact solution, neglecting the transcendentally small term $e^{-1/\varepsilon}$. This is clearly an accident and would again be spoiled by trivial

modifications of the problem. The domains of overlap are, to order unity and ε, respectively:

(4.21c)   for f_0 and g_0: $\mathscr{D}_1 \cap \mathscr{D}_2 = \{\eta | \varepsilon \ll \eta \ll 1\}$;

(4.21d)   for $f_0 + \varepsilon f_1$ and $g_0 + \varepsilon g_1$: $\mathscr{D}_1 \cap \mathscr{D}_3 = \{\eta | \varepsilon \ll \eta \ll \sqrt{\varepsilon}\}$.

4.6. Matching reconsidered. We can now see a posteriori why the matching methods used in the construction earlier worked. Consider first f_0 and g_0. Their overlap domain is given by (4.21c). For any η in this domain,

$$(4.22) \qquad\qquad \lim_\eta (f_0 - g_0) = 0$$

according to the fundamental lemma (3.24). For any limit with η in the overlap domain, $x = \eta x_\eta$ is small and $x^* = \eta x_\eta / \varepsilon$ is large. More specifically,

$$(4.23a) \qquad\qquad \lim_\eta x = 0 \qquad \text{for } \eta \ll 1,$$

$$(4.23b) \qquad\qquad \lim_\eta x^* = \infty \qquad \text{for } \varepsilon \ll \eta.$$

It follows that, for η in the overlap domain,

$$(4.24a) \qquad\qquad \lim_\eta f_0 = f_0(0),$$

$$(4.24b) \qquad\qquad \lim_\eta g_0 = g_0(\infty).$$

Since the numbers given by the right sides above are finite, (4.22) implies (4.9c).
  Similarly,

$$(c) \qquad (f_0 + \varepsilon f_1) - (g_0 + \varepsilon g_1) = bx^2 - \varepsilon 2bx + (1 - a - b)e^{-x^*} + \varepsilon 2be^{-x^*}.$$

With η in the overlap domain (4.21d), $\lim_\eta$ of the exponential terms divided by ε vanishes. Furthermore,

$$(d) \qquad\qquad \lim_\eta [(b\eta^2 x_\eta^2 / \varepsilon) - 2b\eta x_\eta] = 0$$

since $\eta \ll \sqrt{\varepsilon}$. Hence,

$$(4.25) \qquad\qquad \lim_\eta \frac{(f_0 + \varepsilon f_1) - (g_0 + \varepsilon g_1)}{\varepsilon} = 0.$$

This shows that the matching obtained by use of (4.15) was indeed correct.[8]

  We have seen that the formal matching rules (4.9c) and (4.15) may be justified by the application of Lemma (3.24). As stated in the discussion of that lemma matching is by its nature a comparison of two approximations in their domain of overlap.[9] On the basis of this one may of course derive formal rules and recipes for matching in which the concept of overlap does not appear explicitly. Such rules may, however, be of limited validity, as is certainly the case with (4.9c).

---

[8] In subsequent sections we shall see that matching is used not only to fix the values of certain constants but also to determine the function $\beta_j(\varepsilon)$ of (3.10) and (3.11).

[9] Note that for the matching theory it is crucial to use intervals with moving endpoints in discussing the domain of validity of an approximation. If we had used only the fact that f_0 is valid on any interval $[x_0, 1]$, $x_0 = \text{const.} > 0$, then the whole basis for matching would be missing. Also the meaning of such vague statements that f_0 is valid "away from the origin" and g_0 "near the origin" has now been made precise with the aid of the extended concept of interval.

On the other hand, once we know a region of overlap, matching can always be carried out, e.g., by applying Lemma (3.24).

This, however, leads us to one of the central problems in perturbation theory: How can one justify *a priori* that two approximations have a domain of overlap? Some heuristic ideas about this will now be discussed.

4.7. Formal limits of equations and formal domains of validity. In § 3 we have defined η-limits of functions, and also domains of validity. Analogous concepts will now be introduced for equations.

As before we define x_η by $x = \eta x_\eta$. Introducing this variable into (4.1) we obtain

$$(4.26) \qquad \frac{\varepsilon}{\eta^2} \frac{d^2 y}{dx_\eta^2} + \frac{1}{\eta} \frac{dy}{dx_\eta} - a - 2\eta b x_\eta = 0.$$

Each term has a *formal order* in ε determined by the function of ε occurring explicitly. For instance, the first term is formally of order $\varepsilon \eta^{-2}$, since we consider $d^2 y/dx_\eta^2$ to be formally of order unity. Similarly, x_η is considered to be of order unity formally; hence the last term is of order η. Since we may multiply (4.26) by an arbitrary function of ε, it is only the relative order of terms which is important. For each η we now multiply (4.26) by a suitable function of ε such that the highest order terms (or term) are of order unity, and then take the limit of the equation as ε tends to zero. Note that in this limit x_η is fixed; similarly derivatives $d^2 y/dx_\eta^2$ and dy/dx_η are formally unchanged. We are actually applying the limit process $\lim_\eta$ to equation (4.1a). We attach the qualifier "formal" since actually dy/dx_η is a function of both x_η and ε and will change as ε tends to zero; we are not even concerned whether this limit exists. However, the limit must exist wherever functions of ε are indicated explicitly. Thus we must require that η be compatible with unity and with ε; i.e., $1 \cdot \eta^{-1}$ and $\varepsilon \cdot \eta^{-1}$ must tend to definite values, possibly infinite, as ε tends to zero. After the limit has been carried out we may rewrite the resulting equation in x again.

Depending on the o-class of η we then find the following formal limits[10] of (4.1a) (to avoid irrelevant complications we take $\eta = 1$ and $\eta = \varepsilon$ in deriving (4.27b) and (4.27d) respectively):

$$(4.27a) \qquad \operatorname{ord} \eta > 1: \quad -2bx = 0,$$

$$(4.27b) \qquad \operatorname{ord} \eta = \operatorname{ord} 1: \quad dy/dx - a - 2bx = 0,$$

$$(4.27c) \qquad \operatorname{ord} \varepsilon < \operatorname{ord} \eta < \operatorname{ord} 1: \quad dy/dx = 0,$$

$$(4.27d) \qquad \operatorname{ord} \eta = \operatorname{ord} \varepsilon: \quad \varepsilon \frac{d^2 y}{dx^2} + \frac{dy}{dx} = 0,$$

---

[10] This is an oversimplified version of the possible limiting equations. (i) One may introduce a coordinate change of the type (3.7) before carrying out the limits. An example of this is the transformation due to Bush, discussed in § 6. (ii) One may consider different orders of magnitude of the dependent variable. The classical example is the treatment of the stream function in flow at high Reynolds numbers (see, e.g., Van Dyke (1964)). An interesting example is the corner-layer, actually proposed by Kaplun, in the model equation discussed in Cole (1968, p. 32ff.).

An important modification of the idea of limiting equations is discussed after (6.3).

(4.27e) $\operatorname{ord} \eta < \operatorname{ord} \varepsilon: \quad d^2 y/dx^2 = 0.$

Above, the equation on the right is obtained when η is in the function class on the left.

We note that if η is such that $\lim_\eta$ applied to the "full" equation (4.1a) yields (4.27c), then the formal η-limit of (4.27b) also yields (4.27c). In a sense (4.27b) is "rich enough" to contain (4.27c). We make the following definitions.

(4.28) DEFINITION. If E is an equation and $\lim_{\eta_1} E = E_1$, $\lim_{\eta_2} E = E_2$ and also $\lim_{\eta_2} E_1 = E_2$, we say that E_1 contains E_2 (relative to E).

(4.29) DEFINITION. The formal domain of validity of an equation F, relative to the "full" equation E, is the $\operatorname{ord} \eta$ such that $\lim_\eta E$ is either F or an equation contained in F.

According to the definition just made, the validity of (4.27b) is

(4.30a) $\mathscr{D}_1 = \{\eta | \operatorname{ord} \eta > \operatorname{ord} \varepsilon\},$

and that of (4.27d) is

(4.30b) $\mathscr{D}_2 = \{\eta | \operatorname{ord} \eta < \operatorname{ord} 1\}.$

(The notation is the same as in (4.20a) and (4.21a).) One may object that (4.27a) is a nonsensical equation. However, the point is that in the expression

(d) $$\varepsilon \frac{d^2 y}{dx^2} + \frac{dy}{dx} + a - 2bx \equiv \frac{\varepsilon}{\eta^2} \frac{d^2 y}{dx_\eta^2} + \frac{1}{\eta} \frac{dy}{dx_\eta} + a - 2b\eta x_\eta,$$

the last term is of larger order than the other terms provided $1 \ll \eta$.

4.8. Outer and inner limits and equations. Clearly (4.27b) and (4.27d) are the two important equations; one of them or both contain each of the three other equations but neither of them is contained in the other. We shall call them the principal equations. They are obtained by the limit processes indicated above. As representative of $\operatorname{ord} 1$ we take the function $\eta = 1$; the corresponding x_η is then x. Similarly we take ε as representative of $\operatorname{ord} \varepsilon$; the corresponding x_η is x^* as defined by (4.6). We define

(4.31a)   outer limit: $\lim_{\text{out}} F(x, \varepsilon) = \lim F$ as $\varepsilon \downarrow 0$, x fixed;

(4.31b)   inner limit: $\lim_{\text{in}} F(x, \varepsilon) = \lim F$ as $\varepsilon \downarrow 0$, $x^* = x/\varepsilon$ fixed.

We call (4.27b) and (4.27d) the outer and inner equations respectively and x and x^* the outer and inner variables. The outer and inner limits are called principal limits; the $\lim_\eta$ with $\varepsilon \ll \eta \ll 1$ are the intermediate limits. In the present case we spot the principal limits by the fact that they give the two principal equations. The use of (4.4a) and (4.9a) for determining f_0 and g_0 respectively can now be justified by the fact that they are the principal equations. Also, the assumption made earlier that the boundary layer is of thickness ε can now be seen to depend on the fact that the inner limit is principal.

Even within the principal limits there is a hierarchy. The outer limit being in a sense the "first" limit since f_0 can be determined by itself but g_0 must be determined by matching with f_0. In a physical problem this hierarchy is often natural. The solution f_0 represents an unperturbed state. A perturbation of

magnitude ε is introduced; this may mean literally introduced in an experiment or in the theoretical consideration of a small effect which the theory previously had neglected. The effect of the perturbation is represented mathematically by adding the term $\varepsilon(d^2y/dx^2)$ to the equation for the unperturbed state. Typical for a layer-type singular perturbation problem is that while the perturbation is of magnitude ε its effect may be of order unity although only in a very small region. In our case the effect of the perturbation is of order unity near the origin as expressed by $g_0(x^*)$, a fact related to the new boundary condition imposed there which was not imposed on the solution for the state f_0. Thus often purely physical considerations determine which is the first principal equation.[11] The above ideas are illustrated by the discussion of flow at high Reynolds numbers in § 7; in fact (4.1), with $b = 0$, was introduced by Friedrichs (1953) as a model equation for this case.

Returning to our main problem of heuristic justification for our asymptotic construction we see that while the outer and inner equations have been justified, the principle of matching has not been explained. We have already seen that matching of f_0 and g_0 can be explained a priori if we can find an a priori justification for their respective domains of validity. For this we again rely on a study of limits of equations.

4.9. Heuristic principle for domains of validity. As the basis for our discussion we shall use the following heuristic principle due to Kaplun.

(4.32) PRINCIPLE. If y is a solution of an equation E and E^* is an approximate equation, then there exists a solution y^* of E^* whose actual domain of validity (as an approximation to y) includes the formal domain of validity of E^* (as an approximation to E).

In subsequent sections we shall discuss important modifications of this principle. The formulation above is so to speak only a starting point for the discussion. It ceases to be valid when small terms have large integrated effects.[12] In our present example, however, this principle works: The domains of validity of (4.27b) and f_0 on the one hand and the domains of validity of g_0 and (4.27d) on the other hand coincide exactly.

Earlier we justified the asymptotic results by comparison with the exact solution. We now have enough material for an a priori justification. We shall review this, using ideas presented above. We perform the various limits on the equation (4.1a). This leads automatically to the two principal limits and to the two principal equations (4.27b) and (4.27d). We notice that these equations together cover the entire interval $[0, 1]$ (which would not be true of any single

---

[11] Thus the first principal (outer) limit is characterized both by the fact that when applied to the equation it gives a distinguished approximated equation (as described above) and when applied to the solution it gives the unperturbed state. In a reasonable physical theory these two characterizations are equivalent. As will be seen from Examples II and III the inner limit need not correspond to a distinguished equation.

[12] The basic assumption underlying perturbation methods is a sometimes very vague notion that "neighboring equations have neighboring solutions." The heuristic principle (4.32) and modifications thereof are examples of this. We can formalize clearly the idea of solutions being close (as was done in § 3). However, the basic difficulty is to decide when equations are close; this would involve an estimate of integrated effects. These difficulties will be discussed in § 5 and § 6.

equation). Equation (4.27b) is valid for any finite x since its domain includes functions of order unity. Equation (4.27d) is formally valid for all infinitesimal values of x, i.e., it is valid on any interval $[0, \eta]$ provided $\eta \ll 1$. Equation (4.27b) is also formally valid for infinitesimal values of x provided their order of magnitude is greater than ε. The domain of overlap is

$$(4.33) \qquad \mathscr{D}_1 \cap \mathscr{D}_2 = \{\eta | \varepsilon \ll \eta \ll 1\}.$$

We now use the heuristic principle (4.32). It follows that (4.27b) must have a solution f_0 valid in the domain (4.30a). Since this includes the point $x = 1$, as explained above, it must satisfy the boundary condition at $x = 1$. This determines the solution completely to be (cf. (4.5))

$$(4.34) \qquad f_0 = (1 - a - b) + ax + bx^2.$$

Similarly, equation (4.27d) must have a solution g_0 satisfying the boundary condition at $x = 0$. This gives

$$(4.35) \qquad g_0 = C(1 - e^{-x^*}),$$

where the constant cannot be determined by the condition at $x = 1$ since the domain of validity of g_0 cannot be expected a priori to include finite x. However, we have now established the domain of overlap of f_0 and g_0 and may now apply Lemma (3.24). As shown earlier (see (4.22)) this gives the correct matching.

4.10. Higher order terms. The heuristic ideas introduced above may also be used to justify the construction of higher order terms. We write

$$(4.36) \qquad y = f_0 + \varepsilon \tilde{f}_1.$$

Then (4.1a) gives

$$(4.37) \qquad \varepsilon \frac{d\tilde{f}_1}{dx} + \varepsilon \cdot 2b + \varepsilon^2 \frac{d^2\tilde{f}_1}{dx^2} = 0.$$

This is an exact equation. We approximate to $\tilde{f}_1$ by f_1 obeying

$$(4.38a) \qquad df_1/dx + 2b = 0.$$

The domain of validity to order ε is still given by (4.30a). From (4.1c) it follows that the proper boundary condition is

$$(4.38b) \qquad f_1(1) = 0,$$

and hence,

$$(4.39) \qquad f_1 = 2b - 2bx.$$

Similarly we put

$$(4.40) \qquad y = g_0 + \varepsilon \tilde{g}_1$$

which gives

$$(4.41a) \qquad \varepsilon^2 \frac{d\tilde{g}_1}{dx^2} + \varepsilon \frac{d\tilde{g}_1}{dx} - a - 2bx = 0.$$

The exact boundary conditions are

(4.41b) $\tilde{g}_1(0) = 0$,

(4.41c) $\tilde{g}_1(1) = (a + b)/\varepsilon$.

We note, considering the definition (4.40) and the expression for g_0 given by (4.10), that the problem posed by (4.41) is an exact reformulation of the original problem (4.1). We shall now try to attack problem (4.41) directly with the methods previously applied to problem (4.1). To find an inner solution we approximate to $\tilde{g}_1$ by a function g_1 satisfying an equation analogous to (4.27d):

(4.42a) $$\varepsilon^2 \frac{d^2 g_1}{dx^2} + \varepsilon \frac{dg_1}{dx} - a = 0$$

and the boundary condition

(4.42b) $g_1(0) = 0$.

We expect the second boundary condition to be replaced by some matching condition. However, the corresponding outer equation cannot be formed by our previous restricted method of limit processes. The outer limit of (4.41a) does not yield a differential equation. In fact, both terms containing derivatives vanish under any $\lim_\eta$ with $\varepsilon \ll \eta$. On the other hand, if $\eta = \varepsilon$, then $\lim_\eta$ yields the inner equation (4.42a). However, if we assume an outer approximation of the form

(4.43) $\tilde{g}_1 \simeq q_0/\varepsilon + q_1$,

then applications of the outer limit yield

(4.44a) $$\frac{dq_0}{dx} - a - 2bx = 0,$$

(4.44b) $q_0(1) = a + b$,

(4.44c) $$\frac{d^2 q_0}{dx^2} + \frac{dq_1}{dx} = 0,$$

(4.44d) $q_1(1) = 0$

with solutions

(4.45a) $q_0 = ax + bx^2$,

(4.45b) $q_1 = 2b(1 - x)$.

The solution g_1 of (4.42a,b) which matches with (4.45a,b) is then the one given previously by (4.17). The form (4.43) is suggested by the outer boundary condition (4.41c) for $\tilde{g}_1$. Deriving the equation for q_0 amounts to applying the outer limit to (4.41a) keeping $\varepsilon\tilde{g}_1$ fixed; this is necessary because when x is of order unity the dependent variable is of order ε^{-1} rather than of order unity (cf. footnote 10).

We note that by definition $g_0 + \varepsilon g_1$ approximates to y to order ε in the same domain in which g_1 approximates to $\tilde{g}_1$ to order unity. As shown by (4.21b) the actual domain of validity of g_1 is $\mathscr{D}_3 = \{\eta | \eta \ll \varepsilon^{1/2}\}$. On the other hand, the formal

domain of validity of (4.42a) seems to be $\mathscr{D}_2 = \{\eta | \eta \ll 1\}$. This seems to contradict our heuristic principle (4.32). However, reconsidering the basis for the principle, we are looking for the "effect" of the term $-2bx$. In a linear equation this effect is completely independent of the effect of the term $(-a)$. Hence we neglect the third term and compare $(-2bx)$ with the derivatives. For $\varepsilon \ll \eta$ the first derivative dominates and we find that $(-2bx)$ is formally smaller than $\varepsilon(d\tilde{g}_1/dx)$ when $\eta \ll \varepsilon^{1/2}$, i.e., exactly in the domain $\mathscr{D}_3$. The largest part of the effect of $(-2bx)$ is

(a)
$$\frac{1}{\varepsilon} \int_0^x 2bs\, ds = \frac{bx^2}{\varepsilon}.$$

We note that in $\mathscr{D}_3$ this is $o(1)$. Thus in $\mathscr{D}_3$, $\tilde{g}_1$ is never large (if we abstract from the effect of the term $(-a)$).

In a practical computation one would of course never compute an outer approximation to $\tilde{g}_1$ such as is given by (4.43). Our investigation of $\tilde{g}_1$ served only the purpose of showing that the heuristic principle (4.32), suitably interpreted, still applied to the domain of validity of g_1. The functions $\tilde{f}_1$ and f_1 could be discussed in a similar manner. However, no difficulties arise in that case.

The justification for the construction of f_k and g_k for $k > 1$, shown earlier, proceeds in the same way.

As further positive and negative illustration of our heuristic principle we notice that equation (4.27c) has a solution

(4.46a)
$$y \approx l = 1 - a - b$$

which is uniformly valid to order unity in the same domain where (4.27c) is valid. Similarly, (4.27e) has the solution

(4.46b)
$$y \approx m = (1 - a - b)x/\varepsilon$$

which is valid to order unity in the domain where (4.27e) is valid. Neither solution is needed in the asymptotic construction because each is contained in solutions already used. Equation (4.27c) is contained in (4.27b) in the sense of Definition (4.28). We may then also regard (4.27c) as an approximation to (4.27b). By our heuristic principle (4.32), equation (4.27c) has a solution, namely (4.45) which is an approximation to f_0 uniformly valid to order unity in the domain where (4.27c) is a formally valid approximation to (4.27b). In the same sense, the solution (4.45) is an approximation to g_0 in the domain where (4.27c) is an approximation to (4.27d). In this sense the function l of (4.45) is contained in both f_0 and g_0. Similarly, (4.27e) is contained in (4.27d), and accordingly, (4.27e) has a solution, namely m of (4.46b), which is contained in g_0. The function (4.46b) does not match with f_0; it gives the initial slope at $x = 0$ but does not level off to a constant which can be matched with the constant part of f_0 for x small. There is no reason for overlap since the domains of validity of (4.27b) and (4.27e) do not overlap.

4.11. Composite expansion. We have found two expansions which together cover the entire interval $[0, 1]$ although neither is uniformly valid in the entire interval. Thus we have "in principle" obtained a uniformly valid solution. For practical purposes, which is always a prime consideration in applied mathematics, we may want to construct one expansion which is actually uniformly valid in the

whole interval $[0, 1]$:

$$(4.47) \qquad y \simeq \sum \varepsilon^k h_k(x, \varepsilon).$$

Here h_0 should be uniformly valid to order unity in the entire interval $[0, 1]$. This will be achieved if we can construct h_0 in such a way that f_0 and g_0 approximate h_0 to order unity in the same domains (or at least in domains together covering $[0, 1]$) where they approximate the exact solution to order unity. There are several ways of constructing such an h_0. In the present case the most convenient method is the additive one.[13] One puts

$$(4.48) \qquad h_0 = f_0 + g_0 - (\text{common part of } f_0 \text{ and } g_0).$$

The common part is the constant $1 - a - b$. This yields

$$(4.49) \qquad h_0 = (1 - a - b)(1 - e^{-x/\varepsilon}) + ax + bx^2.$$

The justification is that we write

$$(4.50) \qquad y = f_0 + \text{error} + O(\varepsilon).$$

Now f_0 has an error of order unity only for $x = O(\varepsilon)$. Thus to calculate the error to order unity we need an equation for

$$(4.51) \qquad z = y - f_0$$

which is valid in some domain $\eta \ll \eta_0$ where η_0 is some function $\gg \varepsilon$. Using reasoning explained above one finds that a first order approximation to z is

$$(4.52) \qquad z \simeq -(1 - a - b) e^{-x/\varepsilon}$$

which gives (4.49).

4.12. Position of boundary layer. So far we have tacitly assumed that the boundary layer occurs at $x = 0$. Assume instead that it occurs at a point x_0. We then define x_η by $\eta x_\eta = x - x_0$, in particular, x^* by $\varepsilon x^* = x - x_0$. Proceeding as above we are led to a boundary-layer solution $Ae^{-(x-x_0)/\varepsilon} + B$. However, since ε is positive, the exponential will tend to infinity whenever $x < x_0$. Matching is then impossible if $x_0 > 0$. By the same argument we see that if $\varepsilon < 0$, then the boundary layer occurs at $x_0 = 1$.

4.13. Inner and outer expansions and limit-process expansions. We have constructed two asymptotic expansions of y:

$$(4.53a) \qquad y \simeq \sum_{k=0} \varepsilon^k f_k(x),$$

$$(4.53b) \qquad y \simeq \sum_{k=0} \varepsilon^k g_k(x^*).$$

The first expansion, the *outer expansion*, is valid except near the origin; the second one, the *inner expansion*, is valid near the origin. These are the two principal expansions; their domains of validity cover the entire interval $[0, 1]$, and no other principal expansion is needed. More precisely we require the following for outer

---

[13] For other methods, see Van Dyke (1964).

and inner expansions: The domains of validity to order unity of f_0 and g_0 should have a nonempty intersection (overlap domain) and their union should be the entire interval $[0, 1]$. The domains of validity to order ε of $f_0 + \varepsilon f_1$ and $g_0 + \varepsilon g_1$ should have the same property. These requirements are symmetrical for the two expansions. However, as explained earlier, it is f_0 which is being perturbed, and g_0 describes the result in the thin region where the result of the small perturbing term is of order unity. Thus (4.53a) is the first principal expansion, and as a practical consequence, f_0 is determined before g_0.

From the exact solution, or from (4.53) we find

(4.54a) $$\lim_{\text{out}} y = f_0,$$

(4.54b) $$\lim_{\text{out}} \frac{y - f_0}{\varepsilon} = f_1,$$

and we thus obtain $f_0 + \varepsilon f_1 + \cdots$ by a repeated application of the outer limit. Such an expansion will be called an outer-limit expansion. Similarly, a repeated application of the inner limit gives the inner-limit expansion $g_0(x^*) + \varepsilon g_1(x^*) + \cdots$.

In the present example the outer-limit expansion is actually identical with the outer expansion (the analogous statement is true for the inner expansion). They are, however, always conceptually different, and examples in subsequent sections will show that they may be actually different. There is no a priori reason why the inner-limit expansion should match the outer-limit expansion. On the other hand, the method used here of systematically employing limit processes to equations (rather than to functions) and then studying the domains of validity of solutions of these equations gives heuristic reasons why the outer and inner expansions built up of such solutions do match.

4.14. Note on terminology. In Kaplun and Lagerstrom (1967) the notion of characterizing an expansion by its domain of validity was emphasized. The term "intermediate expansion" was used in the sense in which "inner expansion" is used here, whereas the term "outer expansion" was still used to denote a limit process expansion. However, a term must be introduced to denote the expansion corresponding to the outer one but determined by its domain of validity. Furthermore, "intermediate expansion" may wrongly suggest an expansion obtained by repeated application of some intermediate limit. For this reason a deliberate change of terminology has been made in the present paper. The terms "inner" and "outer expansions" refer here to expansions defined by their domains of validity. Limit-process expansions are explicitly labeled as such. While this differs from the terminology presently used in the literature, the change was considered desirable for reasons stated. It also serves to emphasize that expansions defined by their domains of validity are of more basic importance than expansions defined by limit processes.

4.15. Techniques of matching. As discussed above the basis for the matching of two asymptotic expansions is the existence of a domain of overlap. As an actual technique for utilizing this, one may write out

(4.55) $$d_0 = f_0 - g_0,$$

where f_0 is given by (4.5) and g_0 by (4.10a). It is immediately obvious that d_0 can tend to zero under an intermediate limit only if one puts $C = 1 - a - b$. Similarly we define

$$(4.56) \qquad\qquad d_1 = (f_0 + \varepsilon f_1) - (g_0 + \varepsilon g_1)$$

and it is obvious that the constant C_1 in (4.14) must be put equal to $2b$ in order for the η-limit of $d_1 \varepsilon^{-1}$ to tend to zero under some intermediate limit. We see that the technique of the intermediate limit makes precise and justifies the matching rule (4.15) as well as the vague justification of (4.9c) given right after that rule. The connection is that under an intermediate limit x tends to zero and x^* to infinity. It is of course unnecessary to actually use the x_η notation, except for pedagogical purposes or for very tricky cases. Neither is it necessary to use any particular η. As a matter of practical technique one usually does not try to determine a priori the extent of overlap but takes an experimental approach instead. As long as there is some η-region such that the η-limit makes the difference vanish, one assumes that there is overlap. Sometimes, however, it is necessary to exercise a little more caution, as will be illustrated in subsequent sections. Formal success in matching is not always a guarantee of correctness of matching.

Thus the analysis of *why* asymptotic expansions match leads to a very simple technique for actually carrying out the matching.[14]

The technique of applying the intermediate limit to the difference of two partial expansions may be compared with the older[15] technique of interchanging outer and inner limits. For instance, (4.9c) may be written equivalently,

$$(4.57) \qquad\qquad \lim_{\text{out}} \lim_{\text{in}} y = \lim_{\text{in}} \lim_{\text{out}} y.$$

Also applying the inner limit twice (i.e., repeatedly, as illustrated by (4.54)) to $f_0 + \varepsilon f_1$ one finds

$$\text{(a)} \qquad\qquad (1 - a - b) + \varepsilon(ax^* + 2b),$$

whereas applying the outer limit twice to $g_0 + \varepsilon g_1$ (with g_1 given by (4.14)) gives

$$\text{(b)} \qquad\qquad (1 - a - b) + ax + \varepsilon C_1.$$

Requiring (a) and (b) to coincide we find the correct result $C_1 = 2b$.

We see that for the example of the present section the two techniques give the same result. We shall show why this is so by considering a generalization of the present problem.

The general structure of our problem is the following. We have an outer expansion, which at the same time is an outer-limit expansion

$$(4.58a) \qquad\qquad y \simeq \sum_{j=0} \varepsilon^j f_j(x),$$

where each f_j has a Taylor expansion

$$(4.58b) \qquad\qquad f_j(x) = \sum_{k=0}^{\infty} f_{jk} x^k.$$

---

[14] A large number of examples are found in Cole (1968). The basic ideas as well as the technique were introduced by Kaplun in the mid-fifties.

[15] Some references are given in Lagerstrom (1957).

The "error," that is, y minus the outer expansion, consists of terms which die out exponentially for large values of x^*. These terms may be neglected, to any order ε^n, whether we do matching by an intermediate limit (where $\varepsilon \ll \eta$) or by applying the outer limit repeatedly to the inner expansion. Thus the inner expansion (which is also an inner-limit expansion) is, for matching purposes, essentially a rewrite of the outer expansion in inner variables. Whether we use intermediate limits or rewrite in inner variables, we have to use the Taylor expansion (4.58b) of the outer expansion. Thus we have to consider terms of the form

$$(4.59) \qquad T_{jk} = f_{jk}\varepsilon^j x^k = f_{jk}\varepsilon^j \eta^k x_\eta^k = f_{jk}\varepsilon^{j+k} x^{*k}.$$

Consider now matching to order ε^n. Define

$$(4.60a) \qquad A_n = \sum_{j=0}^{n} \varepsilon^j f_j(x) = n + 1 \text{ terms of the outer expansion},$$

$$(4.60b) \qquad B_n = n + 1 \text{ terms of the inner expansion},$$

$$(4.60c) \qquad A_n^* = n + 1 \text{ terms of the inner-limit expansion of } A_n,$$

$$(4.60d) \qquad B_n^* = n + 1 \text{ terms of the outer-limit expansion of } B_n.$$

We see that T_{jk} will occur in B_n if and only if

$$(c) \qquad j + k = m \leqq n.$$

It will then be written

$$(d) \qquad f_{jk}\varepsilon^m x^{*k}.$$

The criterion for its occurrence in A_n^* is the same. In fact, neglecting exponentially small terms, we see that B_n consists entirely of rewrite terms such as (d) and is identical with A_n^*. Conversely, consider a term such as (d) with $m \leqq n$. Rewritten in outer variables it becomes

$$(e) \qquad f_{jk}\varepsilon^{m-k} x^k.$$

Since $0 \leqq m - k \leqq n$, this term occurs in B_n^*. Hence we find the matching criterion

$$(4.61) \qquad A_n^* = B_n^*.$$

Let us instead consider matching by an intermediate limit

$$(4.62) \qquad \lim_\eta \varepsilon^{-n}(A_n - B_n) = 0.$$

The limit is zero by two mechanisms: (i) Certain terms in A_n and B_n must cancel identically when the difference is formed. These are obviously the terms T_{jk} with $j + k \leqq n$. (ii) Other terms tend to zero by themselves, even after division by ε^n. This includes all exponentially small terms of B_n. Furthermore, since the nonexponential terms of B_n come from rewriting A_n, it follows that the domain of validity of A_n extends arbitrarily close to ord ε. By taking ord η sufficiently close to ord ε we see that $\varepsilon^{-n}T_{jk}$ tends to zero under some intermediate limit if and only if $j + k > n$. The most critical case is $j = 0$, $k = n + 1$ which restricts η to be $\ll \varepsilon^{n/n+1}$. The domain of overlap to order ε^n is hence

$$(4.63) \qquad \mathscr{D}_n = \{\eta | \varepsilon \ll \eta \ll \varepsilon^{n/n+1}\}.$$

We thus see that the terms which remain in forming A_n^* and B_n^* are exactly those which must cancel identically in the difference expression in (4.62). The terms that get wiped out in forming A_n^* and B_n^* are exactly those that tend to zero by themselves under a suitable η-limit in (4.62).

Matching techniques will be discussed further at the end of § 6.

5. Second model equation.[16] As Example II we take a function $y(x;\varepsilon)$ defined by the equation

(5.1a)
$$\frac{d^2y}{dx^2} + \frac{n-1}{x}\frac{dy}{dx} + y\frac{dy}{dx} = 0$$

and the boundary conditions

(5.1b) $y = 0$ at $x = \varepsilon$;

(5.1c) $y = 1$ at $x = \infty$.

The cases $n = 1, 2, 3$ will be discussed in detail.

We shall give (5.1a) a somewhat vague but very convenient interpretation. The first two terms represent the Laplacian of a spherically symmetric function in n dimensions, the variable x being the radial variable. We shall think of y as temperature; then (5.1a) is the equation for the equilibrium distribution of temperature with the term $y(dy/dx)$ representing some fictitious nonlinear heat loss. The boundary conditions state that the temperature is equal to one at infinity and zero at the surface of a small n-sphere of radius ε. We interpret this as a perturbation problem in the following way: In the absence of the sphere the temperature is equal to one everywhere. This situation is perturbed by the introduction of a cool ($y = 0$) sphere. Since its radius is $o(1)$, we expect the perturbation to be small (for $n = 2, 3, \cdots$) except near the surface of the sphere where the boundary condition forces the temperature to be zero. Similarly, if the radius ε shrinks to zero and the sphere disappears, the temperature will revert everywhere to one. More precisely, if one observes the temperature at a fixed point (in terms of x), the influence of the sphere will vanish as ε tends to zero; i.e., y will tend to one. This convergence is, however, nonuniform near the sphere. An observer whose x-coordinate varies as εC, $C = $ const., is expected to observe a strong influence of the sphere; in particular, if $C = 1$, he always observes $y = 0$. We thus have a singular perturbation problem. The heuristic argument suggests

---

[16] The problem discussed here was originally introduced as a model for the asymptotic theory of incompressible flow at low Reynolds numbers as given in Kaplun and Lagerstrom (1957) and Kaplun (1957). Equation (5.1a) corresponds to the Navier–Stokes equations (see § 7) for viscous flow in n dimensions, (5.6b) to the Stokes equations and (5.11) to the Oseen equations in three dimensions and (A.1a) to the Oseen equations in $m - 1$ dimensions. Previously the Stokes equations had been considered as a linearization of the Navier–Stokes equations. Kaplun's analysis of the asymptotic nature of the equations showed that the linearity was a coincidence and would no longer occur for compressible flow (Kaplun (1957)). The asymptotic derivation of the nonlinear Stokes equation for compressible flow is given in Lagerstrom (1964). The model equation of § 6 was introduced to illustrate certain mathematical complications occurring in the compressible case. While the "Stokes equation" (5.6b) is linear for the present example, the corresponding equation in § 6, namely (6.2b), is nonlinear. Much of the material in § 5 and § 6 is based on Lagerstrom (1961). For further discussion of the same model equations, see Bush (1971) and the references given there.

the following principal variables:

(5.2) $x =$ outer variable, $x^* = x/\varepsilon =$ inner variable.

The convergence to one as described above can now be described as

(5.3) $f_0 = \lim_{\text{out}} y = f_0 \equiv 1.$

The heuristic interpretation also suggests that the convergence to unity is uniform in any interval $[x_0, \infty], x_0 > 0$. Note that the terminology agrees with the conventions discussed after (4.31). The outer limit is the first principal limit since it is the state $y \equiv f_0$ which is perturbed by the introduction of a small sphere.

As a negative confirmation of the above argument we consider the case $n = 1$. The one-sphere is an interval of length 2ε. We may think of the one-sphere as two planes perpendicular to the x-axis located at a distance $\pm\varepsilon$ from $x = 0$ in three-dimensional space. Clearly as ε tends to zero the "sphere" does not disappear. The planes decompose space into two independent regions. Physically, the magnitude of ε is irrelevant; it is taken care of by the shift in coordinate $x \rightarrow x - \varepsilon$. In fact, the exact solution is

(5.4) $y(x, \varepsilon) = \tanh\left(\dfrac{x - \varepsilon}{2}\right) = \dfrac{1 - e^{-(x-\varepsilon)}}{1 + e^{-(x-\varepsilon)}}.$

We see that while $f_0 = 1$ is an approximation valid for x large, it is not valid for x of order unity. The function (5.4) has indeed a regular expansion for ε small.

Returning now to the cases $n = 2, 3, \cdots$ we shall see whether a formal study of the limiting equations justifies our choice of variables. As in Example I we introduce the variable x_η by $\eta x_\eta = x$ and rewrite (5.1a) as

(5.5) $\dfrac{d^2 y}{dx_\eta^2} + \dfrac{n-1}{x_\eta} \dfrac{dy}{dx_\eta} + \eta y \dfrac{dy}{dx} = 0$

and obtain various approximating equations valid in different domains of the functions $\eta(\varepsilon)$:

(5.6a) $y\dfrac{dy}{dx} = 0 \quad \text{for } \eta \gg 1,$

(5.6b) $\dfrac{d^2 y}{dx^2} + \dfrac{n-1}{x} \dfrac{dy}{dx} = 0 \quad \text{for } \eta \ll 1.$

For η of order unity the full equation (5.1a) is obtained. Thus, as it should according to the discussion after (4.31), the first principal limit gives a distinguished equation. The inner limit (x^* fixed) gives, however, the same equation as any η-limit with $\eta = o(1)$. Its importance comes rather from the boundary conditions, as discussed earlier. Actually, the inner limit will play a less important role here than in Example I (its role will be even less important in Example III). What matters is that the formal domain of validity of (5.6b) includes all η, such that $\eta = o(1)$, in particular $\eta = \varepsilon$. In accordance with the heuristic principle (4.32), we expect that there exists a solution g_0 of (5.6b) valid to order unity in the same domain. In particular, such a solution must be valid at $x = \varepsilon$, i.e., it must satisfy

the inner boundary condition (5.1b). It will turn out that such a solution is conveniently expressed with the aid of the inner variable. This is natural since the inner boundary condition is given at $x^* = 1$. However, we shall see later that the ε-dependence of the terms in the inner expansion will not always have the simple form: (function of ε) times (function of x^*).

The solution g_0 just described is determined within a constant of integration B (which may depend on ε) and is

(5.7a) $$g_0 = B \ln x^* \qquad \text{for } n = 2,$$

(5.7b) $$g_0 = B(1 - (x^*)^{2-n}) \quad \text{for } n > 2.$$

The constant B must be determined by matching[17] with the outer solution. The function f_0 given by (5.3) seems to be a candidate for the outer solution. It certainly is uniquely determined as the solution of the first order equation (5.6a) since it satisfies the outer boundary conditions. However, the regions of validity of (5.6a) and (5.6b) do not overlap. Assume, however, that the domain of validity of f_0 also includes the η such that ord $\eta = $ ord 1. Then we may use the extension theorem, Theorem (3.25), to prove that its domain of validity also includes the region $\eta \gg \eta_0$ for some $\eta_0 \ll 1$. The domains of validity of f_0 and g_0 then overlap, and matching is possible.

To justify that f_0 is valid for x of order unity we cannot rely solely on the equations. The corresponding equation is the full equation (5.1a). While f_0 is a solution of this equation satisfying the outer boundary condition, it is not uniquely determined since (5.1a) is a second order equation. However, we may instead use our earlier heuristic reasoning which showed that f_0 is the outer limit of y, valid on any interval $[x_0, \infty)$ with $x_0 > 0$, which means that its domain of validity includes the domain $\{\eta | \text{ord } \eta \geqq \text{ord } 1\}$, and that hence the extension theorem may be used. Thus in the present example it is essential that f_0 be an outer limit; it is the combination of this fact and the extension theorem which shows that it has a large enough domain of validity. On the other hand, the domain of validity of g_0 is obtained from a consideration of the validity of the equation which it satisfies. It may happen to be an inner limit, but this is not essential.

Parenthetically we shall show that some special arguments about the domain of validity of f_0 given above are necessary. Its extended domain of validity, and the existence of an overlap domain, cannot be automatically deduced from the "adjacent" but not overlapping domains of validity of (5.6a) and (5.6b). In Example I the function $l(x, \varepsilon) = 1 - a - b$ satisfying (4.27c) is valid for ord ε < ord η < ord 1, and the function $m(x, \varepsilon) = (1 - a - b)x/\varepsilon$ satisfying (4.27e) is valid for ord η < ord ε. However, these two functions do not match in spite of the fact that $l(x, \varepsilon)$ satisfies (4.27d). Instead, each of the functions match with a nonconstant solution of (4.27d), namely $g_0(x^*)$ as given by (4.10). Note that matching of $m(x, \varepsilon)$, multiplied by an undetermined constant, with $l(x, \varepsilon)$ is

---

[17] In the corresponding problem in fluid dynamics the fact that (5.7a) cannot satisfy the boundary condition at infinity (5.1c) was known as the Stokes paradox. It was resolved when Kaplun pointed out that the perturbation problem is singular and that the Stokes solution (5.7) is an inner solution which must satisfy a matching condition but not necessarily the boundary condition at infinity. An elementary presentation of Kaplun's argument is given in Lagerstrom and Cole (1955, p. 873ff.).

formally impossible. In Example II, f_0 can be matched formally with g_0. This, however, is a necessary but not a sufficient condition. As will be illustrated by Example III, formal success of matching is not a sufficient proof of its correctness.

We shall now work out some details for the cases $n = 3$ and $n = 2$.

5.1. The case $n = 3$. The solution of the inner equation (5.6b) is

$$(5.8) \qquad g_0 = 1 - 1/x^*.$$

Here the constant B of (5.6b) has been evaluated to $B = 1$ by matching with $f_0 = 1$. Obviously g_0 tends to f_0 when x^* tends to infinity, i.e., under any η-limit such that $\varepsilon \ll \eta$. Now putting

$$(5.9) \qquad y = 1 + \beta \tilde{f}_1, \qquad \tilde{f}_1 = f_1 + O(\beta),$$

where $\beta(\varepsilon) \ll 1$, we find the exact equation for $\tilde{f}_1$ to be

$$(5.10) \qquad \frac{d^2\tilde{f}_1}{dx^2} + \frac{2}{x}\frac{d\tilde{f}_1}{dx} + \frac{d\tilde{f}_1}{dx} + \beta\tilde{f}_1\frac{d\tilde{f}_1}{dx} = 0.$$

By applying limit processes to this equation we find equations for f_1:

$$(5.11) \qquad \frac{d^2f_1}{dx^2} + \frac{2}{x}\frac{df_1}{dx} + \frac{df_1}{dx} = 0, \qquad \text{ord } \eta = \text{ord } 1,$$

and

$$(5.12a) \qquad \frac{df_1}{dx} = 0 \quad \text{for } \eta \gg 1,$$

$$(5.12b) \qquad \frac{d^2f_1}{dx^2} + \frac{2}{x}\frac{df_1}{dx} = 0 \quad \text{for } \eta \ll 1.$$

Since (5.12a) and (5.12b) are contained in (5.11), this equation is formally valid everywhere, i.e., in the whole interval $[\varepsilon, \infty)$.

The solutions of (5.11) satisfying $f_1(\infty) = 0$ have the form

$$(a) \qquad f_1 = A_1 E_2(x)$$

when E_2 is defined as in the Appendix. To determine A_1 and β by matching we apply an intermediate limit to the difference[18]

$$(b) \qquad d_0 = (1 + \beta f_1) - g_0.$$

Since x tends to zero and x^* to infinity under the intermediate limit, we need the expansion of $1 + \beta f_1$ for x small and the expansion of g_0 for x^* large. The function g_0 is already given in expanded form. From the Appendix we find

$$(c) \qquad 1 + A_1\beta E_2 = 1 + A_1\beta(1/x + \ln x + O(1)).$$

Thus, if we put

$$(d) \qquad A_1 = -1, \qquad \beta = \varepsilon,$$

---

[18] Here we determine the second term of the outer expansion by matching with g_0 alone. The more general rule is that the second terms of both the inner and outer expansions have to be determined by simultaneous matching. See Kaplun and Lagerstrom (1957, p. 590) and Kaplun (1957, p. 598).

then

(e) $$d_0 = -\varepsilon \ln x + O(\varepsilon),$$

which tends to zero in the entire x-range (although not uniformly, it does not vanish for $x \sim e^{1/\varepsilon}$). Hence,

(5.13) $$f_0 + \beta f_1 = 1 - \varepsilon E_2(x).$$

In § 4 we emphasized the use of matching to determine constants of integration not given by boundary conditions. Actually matching is also used to determine the functions $\beta_j(\varepsilon)$ of the expansions (3.10) and (3.11). In § 4 this was trivial but the present section contains several nontrivial examples of determining the β_j.

5.2. Higher order terms. If we now put

(5.14) $$y = g_0(x^*) + \varepsilon \tilde{g}_1(x^*, \varepsilon), \qquad \tilde{g}_1 = g_1(x^*) + o(1),$$

then

(5.15) $$\frac{d^2 \tilde{g}_1}{dx^{*2}} + \frac{2}{x^*}\frac{d\tilde{g}_1}{dx^*} + g_0\frac{dg_0}{dx^*} + \varepsilon\left(g_0\frac{d\tilde{g}_1}{dx^*} + \tilde{g}_1\frac{dg_0}{dx^*}\right) + \varepsilon^2\tilde{g}_1\frac{d\tilde{g}_1}{dx^*} = 0.$$

Neglecting terms of formal order ε and using the value of g_0 given by (5.8), we have

(5.16) $$\frac{d^2 g_1}{dx^{*2}} + \frac{2}{x^*}\frac{dg_1}{dx^*} = \left(\frac{1}{x^*} - 1\right)\left(\frac{1}{x^*}\right)^2.$$

The neglected terms remain of formal order ε if one introduces any x_η, where $\eta x_\eta = x = \varepsilon x^*$, provided $\eta \ll 1$. Hence (5.16) is formally valid in this domain.

The solutions satisfying the inner boundary condition $g_1 = 0$ at $x^* = 1$ have the form

(5.17) $$g_1 = -\frac{\ln x^*}{x^*} - \ln x^* + B_1 g_0.$$

Hence,

(5.18) $$g_0 + \varepsilon g_1 = 1 - \frac{\varepsilon}{x} + \varepsilon\left(\frac{-\varepsilon \ln x + \varepsilon \ln \varepsilon}{x} - \ln x + \ln \varepsilon\right) + \varepsilon B_1\left(1 - \frac{\varepsilon}{x}\right).$$

We expect this to be valid to order ε. The formal domains of validity of the equations for f_1 and g_1 indicate that in matching by taking an intermediate η-limit we must have $\eta \ll 1$. Thus we need to find the expression for f_1 for x small. From the Appendix we find

(5.19) $$f_0 + \varepsilon f_1 = 1 + \varepsilon(-1/x - \ln x - (\gamma - 1) + x/2 + O(x^2)).$$

The difference between (5.19) and (5.18), divided by ε, should tend to zero with ε. However, our construction fails here. The term $\varepsilon \ln \varepsilon$ in (5.18) cannot be matched by anything in (5.19). No choice of the constant B_1 (if considered independent of ε) or of an η-limit can remedy this situation; neither can the unmatched term be eliminated by continuing the expansions to higher order.

The formal remedy is to backtrack to lower order terms. If we insert into (5.18) a term $\zeta(\varepsilon)g_s(x^*)$, then g_s will obey the homogeneous inner equation (5.6b) provided $\varepsilon \ll \zeta$. The obvious choice is

$$(5.20) \qquad \zeta g_s(x^*) = -\varepsilon \ln \varepsilon g_0(x^*).$$

Then $\varepsilon \ln \varepsilon$ disappears from the expansion

$$(5.21) \qquad y \simeq g_0(x^*) - \varepsilon \ln \varepsilon g_0(x^*) + \varepsilon g_1(x^*).$$

Putting $B_1 = 1 - \gamma$ in order to match the terms ε times const., we now find that the difference d_1 between (5.19) and (5.21) is

$$(5.22) \qquad d_1 = \frac{\varepsilon^2 \ln x}{x} - \frac{\varepsilon^2(\gamma - 1)}{x} - \frac{2\varepsilon^2 \ln \varepsilon}{x} + \varepsilon O(x).$$

Matching is now formally possible since

$$(5.23) \qquad \lim_\eta \frac{d_1}{\varepsilon} = 0 \quad \text{for } |\varepsilon \ln \varepsilon| \ll \eta \ll 1.$$

Note that the overlap domain is smaller than anticipated.

While we have succeeded in overcoming our difficulties it is essential for understanding perturbation theory to understand why our a priori estimates fail. We have here run into the phenomena of integrated effects and resulting switchback, two terms which will be explained below after some additional formal calculations.

5.3. Calculation of f_2. If one continues the outer expansion formally as

$$(5.24) \qquad y \simeq 1 + \varepsilon f_1 + \varepsilon^2 f_2,$$

then f_2 obeys (A.3) of the Appendix whose solution is

$$(5.25a) \qquad f_2 = F_2 + A_2 E_2.$$

For small values of x (cf. Appendix),

$$(5.25b) \qquad f_2 = \left(-\frac{\ln x}{x} - \frac{\gamma + 2}{x} - 3 \ln x + O(1) \right) + (2\gamma + 1)\left(\frac{1}{x} + \ln x + O(1) \right),$$

where the constant A_2 has been evaluated to be $2\gamma + 1$.

5.4. Switchbacks, integrated effects. Nothing in our formal expansion procedure indicated that an intermediate term of order $\varepsilon \ln \varepsilon$ was needed until it was found that (5.18) and (5.19) did not match. Terms such as (5.20) whose existence is deduced from inspection of formally higher order terms, such as (5.17), are called switchback terms. The switchback phenomenon is quite common in singular perturbation constructions.[19]

In § 2 we saw an example of how the effect of a term may be of larger magnitude than the order formally indicated. The term of order ε in (2.7a) has,

---

[19] Classical examples from fluid mechanics are the so-called Filon terms discussed by Chang (1961, p. 851ff.) and the $R^2 \ln R$ term in Proudman and Pearson (1957, p. 255). An artificial but very intuitive example is given in Kaplun (1967, p. 15).

over a time interval sufficiently long, an effect of order unity. In our present problem we may write the basic equation as

$$(5.26) \qquad \frac{d^2 y}{dx^{*2}} + \frac{2}{x^*} \frac{dy}{dx^*} = -\varepsilon y \frac{dy}{dx^*}.$$

Suppose that we impose the condition $y = 0$ at $x^* = 1$ and (instead of the boundary condition at $x = \infty$) the condition that, for x^* large, y should match, to appropriate order, with the outer expansion which is considered known. Assume that we have the correct solution for y, $y = f(x^*, \varepsilon)$, and insert its value into the right-hand side, keeping y on the left-hand side as unknown. This second problem is a linear equation with known forcing term $-\varepsilon f(df/dx^*)$, and its solution is of course the same as that of the first problem. The solution of the problem with $\varepsilon \neq 0$ minus the solution of the problem with $\varepsilon = 0$ (which is $g_0(x^*)$) is the integrated effect of the right-hand side. The question is whether this effect is of order ε or of a different order. In our case we hope that if we use $g_0(x^*)$ in the right-hand side, rather than the correct $f(x^*, \varepsilon)$, the answer will change by $O(\varepsilon)$. We deal with a problem of the type: a second order linear operator operating on y equals $\varepsilon \cdot$ (function of x^*) with a boundary condition at $x^* = 1$ and a matching condition for x^* large. Suppose we have no matching condition but a second boundary condition at a point independent of ε, say at $x^* = \infty$, and that the appropriate Green's function exists. The integrated effect of the forcing function is then formally the customary integral involving the forcing function and the Green's function. Now, *either* the integral does not exist, i.e., the problem has no solution, *or* its value is $\varepsilon \cdot$ (function of x^*); i.e., the integrated effect is of order ε (if we include the degenerate case of the result being identically zero). Thus, there is nothing surprising about the magnitude of integrated effects in any reasonable physical problem where we solve a boundary value problem.

However, let us now return to the case in which the outer condition is a matching condition. Let us consider our special case (5.26) with y replaced by $g_0(x^*)$ in the right-hand side. The general solution satisfying $y = 0$ at $x^* = 1$ is

$$(5.27) \qquad \begin{aligned} y &= C\left(1 - \frac{1}{x^*}\right) - \frac{\varepsilon \ln x^*}{x^*} - \varepsilon \ln x^* \\ &= C\left(1 - \frac{\varepsilon}{x}\right) - \frac{\varepsilon^2(\ln x - \ln \varepsilon)}{x} - \varepsilon \ln x + \varepsilon \ln \varepsilon. \end{aligned}$$

There is, of course, no a priori reason to assume C independent of ε. Matching (5.27) with (5.19) to order ε we find that the term $\varepsilon \ln \varepsilon$ is too large and must be gotten rid of with the aid of C. In fact, we must have

$$(5.28) \qquad C = 1 - \varepsilon \ln \varepsilon + \varepsilon(1 - \gamma),$$

which of course agrees with (5.21).

Essential for integrated effects is that in a matching problem we consider intervals with moving endpoints and hence containing different orders of magnitude of x^* (and x). Without the term $-\varepsilon \ln \varepsilon$ the solution would be too large in the matching region (where x^* is very large). But introducing this term

makes it larger than anticipated where x^* is of order unity. Since often a quantity such as dy/dx^* is of physical interest this result may be important.[20]

A systematic justification of a perturbation method must involve a method of estimating integrated effects. Large integrated effects occur especially when an infinite domain is considered or when singularities such as x^{-1} occur in the approximating functions. In the simple case of Example I it is clear that the integrated effects will be of the order formally indicated.

5.5. The case $n = 2$. In this case

$$(5.29) \qquad g_0 = B(\ln x - \ln \varepsilon) = B \ln x^*$$

with $\varepsilon x^* = x$. The difference between g_0 and f_0 is

(a) $\qquad d_0 = B(\ln x_\eta + \ln \eta - \ln \varepsilon) - 1.$

If

$$(5.30) \qquad B = -1/\ln \varepsilon + o(-1/\ln \varepsilon),$$

then

$$(5.31) \qquad \lim_\eta d_0 = 0 \quad \text{for } |\ln \eta| \ll |\ln \varepsilon|.$$

In the present example we see that the overlap domain is very small. In particular,

$$(5.32) \qquad \lim_\eta g_0 = 1 - a \quad \text{for } \eta = \varepsilon^a,$$

so that the functions ε^a, $a > 0$, do not belong to the overlap domain. On the other hand, we note that the outer limit gives the correct result. Hence the overlap domain includes ord 1 and, by the extension theorem, some $\eta \gg 1$.[21]

5.6. Higher order terms. To obtain higher order approximations we proceed as in the case $n = 3$. We assume expansions

$$(5.33a) \qquad y = 1 + \alpha \tilde{f}_1, \qquad \tilde{f}_1 = f_1(x) + o(1),$$

$$(5.33b) \qquad y = g_0(x^*) + \alpha \tilde{g}_1, \qquad \tilde{g}_1 = g_1(x^*) + o(1),$$

where

$$(5.33c) \qquad \alpha = -1/(\ln \varepsilon + b), \qquad b = O(1).$$

We note that

$$(5.34) \qquad e^{-1/\alpha} = \varepsilon e^b.$$

Hence the functions ε^a, $a > 0$, are transcendentally small relative to the α^n. In the exact equation for $\tilde{g}_1$ the highest order terms are

(b) $\qquad \dfrac{\alpha d^2 \tilde{g}_1}{dx^{*2}} + \dfrac{\alpha}{x^*} \dfrac{d\tilde{g}_1}{dx^*} = O\left(\dfrac{\alpha}{x^{*2}}\right).$

---

[20] A concrete physical example is the $R^2 \ln R$ in the drag formula for low Reynolds number flow past a sphere, given in Proudman and Pearson (1967).
[21] This also follows from $|\ln \eta^{-1}| = |\ln \eta|$.

For $x = O(\eta)$ the first term is $O(\alpha \varepsilon^2/\eta^2)$ and the second term is $O(\varepsilon^2/\eta)$. The ratio of the second term to the first is then $O(\eta/\alpha)$. For $n \ll \varepsilon^a$, with $a > 0$, this ratio is transcendentally small. For this reason g_1 will obey the same homogeneous equation as g_0. A similar reasoning will hold for g_2. Matching of $g_0(x^*) + \alpha g_1(x^*) + \alpha^2 g_2(x^*)$ and of $1 + \alpha f_1(x) + \alpha^2 f_2(x)$ is formally possible so that there is no evidence of large integrated effects.

The function f_1 obeys (A.la) with $m = 1$ (see Appendix) and the solution is

$$(5.35) \qquad\qquad f_1 = -A_1 E(x).$$

We write the solution for g_1 as

$$(5.36) \qquad\qquad g_1 = B_1 \alpha \ln x^*.$$

Then, for x small,

(c)
$$\begin{aligned}
d_1 &= f_0 + \alpha f_1 - g_0 - \alpha g_1 \\
&= 1 + \alpha A_1(\ln x + \gamma + O(x)) - B(\ln x - \ln \varepsilon) - B_1 \alpha^2(\ln x - \ln \varepsilon).
\end{aligned}$$

Putting

$$(5.37) \qquad\qquad B = \alpha \quad \text{and} \quad A_1 = 1$$

we find

(d)
$$\begin{aligned}
d_1 &= 1 - \frac{\ln \varepsilon}{\ln \varepsilon + b} - \frac{\gamma}{\ln \varepsilon + b} + \frac{B_1 \ln \varepsilon}{(\ln \varepsilon + b)^2} - B_1 \frac{\ln x}{(\ln \varepsilon + b)^2} + O(x)\alpha \\
&= \frac{b}{\ln \varepsilon} - \frac{\gamma}{\ln \varepsilon} + \frac{B_1}{\ln \varepsilon} - \frac{B_1 \ln x}{(\ln \varepsilon)^2} + O(x)\alpha.
\end{aligned}$$

The obvious choice is to put

$$(5.38a) \qquad\qquad B_1 = 0 \quad \text{and} \quad b = \gamma,$$

i.e.,

$$(5.38b) \qquad\qquad \alpha = -1/(\ln \varepsilon + \gamma).$$

Then

(e)
$$d_1 = O(x)\alpha$$

and

$$(5.39) \qquad\qquad \lim_\eta d_1/\alpha = 0 \quad \text{for } \eta = o(1).$$

If instead we put $b = 0$, then we must have $B_1 = \gamma$ and the term $B_1 \ln x/(\ln \varepsilon)^2$, divided by α, goes to zero only in the very restricted domain where $|\ln \eta| \ll |\ln \varepsilon|$.

The term f_2 obeys (A.3) of the Appendix; hence,

(5.40)
$$\begin{aligned}
f_2 &= 2E(2x) - (e^{-x} \cdot E(x)) + A_2 E(x) \\
&= -(1 + A_2) \ln x - 2 \ln 2 - \gamma - A_2 \gamma - x \ln x + O(x).
\end{aligned}$$

This solution has to be matched with a solution valid near the body

$$(5.41) \qquad\qquad \alpha^2 g_2 = B_2 \alpha^3 \ln x^*.$$

The difference

(f) $$d_2 = (f_0 + \alpha f_1 + \alpha^2 f_2) - (g_0 + \alpha g_1 + \alpha^2 g_2)$$

becomes, with

(5.42) $$A_2 = -1 \quad \text{and} \quad B_2 = -2 \ln 2,$$

(g) $\quad d_2 = d_1 + \alpha^2[-2 \ln 2 - x \ln x + 2 \ln 2 \cdot \alpha(\ln x - \ln \varepsilon) + O(x)].$

Since $d_1 = O(x) \cdot \alpha$,

(h) $$\lim_\eta d_1/\alpha^2 = 0 \quad \text{for } \eta \ll \alpha.$$

Thus,

(5.43) $$\lim_\eta d_2/\alpha^2 = 0 \quad \text{for } |\ln \eta| \ll |\ln \varepsilon| \quad \text{and} \quad \eta \ll \alpha.$$

An essential feature of the discussion of the case $n = 2$ is the curious metamorphosis of the logarithmic function into a constant; for instance, $\ln x^*/-\ln \varepsilon$ tends to unity for $x = O(\eta)$, where η is in the overlap domain. As we saw earlier in the discussion of integrated effects, an essential feature of the analysis of layer-type problems is that we consider regions of x of different order and that the order of a function of x depends on the order of x. Thus $\alpha \ln x^*$ is $O(\alpha)$ for $x = O(\varepsilon)$ but $O(1)$ for $x = O(\varepsilon^a)$, $0 < a < 1$.

In general, according to (5.32), g_0 tends to the constant value $1 - a$ if x decreases as ε^a. In a sense, the logarithmic function g_0 generates all the constants in the interval $[0, 1]$. It was Kaplun's observation (in the corresponding problem in fluid dynamics) that while $\lim_\eta y$ for $\eta \ll 1$ gives a continuum of limits, the same limit processes applied to the equation yield only one equation, namely (5.6b) with $n = 2$. His idea was then to look for a *solution* of this equation which need not be a limit but which would contain all these limits. This solution is in fact $g_0(x^*)$.

We have just discussed in what sense the inner solution senses the y-value ("temperature") at infinity. Actually, of course, it should be matched with a somewhat lower value since $y < 1$ for any finite x. The constant term in the expansion of f_1 gives the correction to order α of the temperature at the apparent infinity. This was taken care of by a special trick of defining α. Similarly $\alpha^2 f_2$ gives a correction to order α^2 of the temperature at the apparent infinity; the inner term $\alpha^2 g_2$ has to agree with this value in the same sense in which g_0 agreed with the value one. Note that the term $y(dy/dx)$ will be transcendentally small in the inner region in the sense that it will never enter the equations for the g_j. However, it will influence the g_j indirectly through the matching conditions.

We shall now make various comments on the results obtained for $n = 2$ and $n = 3$.

5.7. Outer and inner expansions versus limit-process expansions. Since partial sums of the inner and outer expansions are together uniformly valid to an order depending on the number of terms considered we may apply limit processes to these expansions instead of to y. For $n = 3$ we find that the outer and inner limits

of y are f_0 and g_0, respectively. For a repeated application of the inner limit we find

(5.44)
$$\lim_{\text{in}} \frac{y - g_0}{\zeta} = g_s(x^*) = g_0(x^*)$$

with $\zeta = -\varepsilon \ln \varepsilon$. Since

(5.45)
$$\lim_{\text{out}} \frac{y - f_0}{\zeta} = 0,$$

we may put $f_s = 0$. The first two terms of the outer-limit expansion and the inner-limit expansion respectively are then $g_0 + \zeta g_s$ and $f_0 + \zeta f_s$, respectively. Earlier in this section we saw that regarding f_0 as an outer limit helped determine its domain of validity. However, in general a limit-process expansion does not give any indications of the domain of validity. In fact,

(5.46) $(f_0 + \zeta f_s) - (g_0 + \zeta g_s) = 1 - (1 - \varepsilon/x) + \varepsilon \ln \varepsilon(1 - \varepsilon/x).$

This difference, divided by $\zeta = -\varepsilon \ln \varepsilon$, does not tend to zero.

The functions f_0 and g_0 give uniform validity to order unity, i.e., the union of their domains of validity is the entire interval $\varepsilon \leqq x < \infty$ (cf. (3.20b)) and are thus the leading terms of the outer and inner expansions respectively. The term $\zeta g_s(x^*)$ gives higher order validity near $x^* = 1$ and thus gives a correction to g_0 (if one, for instance, is interested in computing $(dy/dx^*)_{x^*=1}$). However, as seen by (5.46), the terms ζg_s and ζf_s do not together give uniform validity to an order higher than unity; in other words, adding these terms does not improve the uniform validity. Actually, the second term of the outer expansion is εf_1. We forget about f_s. It is identically zero; it is only if one insists on limit-process expansions that one needs it for bookkeeping purposes. As the second term of the inner expansion one should take $\varepsilon g_a(x, \varepsilon) = -\varepsilon \ln \varepsilon g_0(x^*) + \varepsilon g_1(x^*)$. If this term is added to g_0 and εf_1 to f_0, then the inner and outer expansions together are uniformly valid to order ε. Thus the second term of the inner expansion does not have the form $\beta(\varepsilon)g(x^*)$. However, if we liberate ourselves from the notion that the inner expansion should be a limit process expansion, we see that there is no reason why it should have such a special form. This point will be discussed further at the end of § 6.

For $n = 2$ we find as usual that the outer limit of y equals one and

(5.47) $\lim_{\text{in}} y = 0, \qquad \lim_{\text{in}} (y - 0)/\alpha = \ln x^*.$

The first term of the inner-limit expansion is then zero. By repeated application of the inner limit we find that the second term of the inner-limit expansion agrees with the first term of the inner expansion. We see thus that inner and outer limits do not match. A priori we can only expect the following for the domains of validity of limits: Let f be the η_0-limit of a function $z(x, \varepsilon)$. By definition $\lim_{\eta_0} (z - f)$ is zero, and the domain of validity of f as an approximation to z must include the functions η_0. By the extension theorem we know that there must exist functions η_1 and η_2, $\eta_1 \ll \eta_0 \ll \eta_2$, such that

(5.48a) $\mathscr{D}' = \{\eta | \eta_1 \ll \eta \ll \eta_2\} = (\text{ord } \eta_1, \text{ord } \eta_2)$

is included in the domain of validity. Now let $\zeta_0 \ll \eta_0$ and let the ζ_0-limit of z be g. Then g is valid in some open order-interval containing

(5.48b) $$\mathscr{D}'' = (\text{ord } \zeta_1, \text{ord } \zeta_2)$$

for some functions $\zeta_1 \ll \zeta_0 \ll \zeta_2$. However, there is nothing in the extension theorem which states that $\mathscr{D}'$ and $\mathscr{D}''$ must overlap. The fact that f_0 was an outer limit determined that it had a domain of validity of the form (5.48a) with $\eta = 1$. However, this trick was useful only once; g_0 had to be determined, not as a limit, but as a solution whose formal domain of validity overlapped with the domain of validity of f_0.

We see that, at least to the order computed here, the same terms appear in the inner and the inner-limit expansions. However, we notice that the first term in the inner-limit expansion is 0 and the second term is $\alpha \ln x^*$. The first term in the inner expansion, however, is $\alpha \ln x^*$. This term, together with the first term in the outer expression, namely $f_0 \equiv 1$, gives uniform validity to order unity.

5.8. Inner limit of the outer expansion. Since the equation for f_1 contains the equation for g_0, we would expect that a solution of the former equation would contain g_0. Actually, for $n = 3$ the inner limit of $1 + \varepsilon f_1$ (which obeys the same equation as f_1) is g_0 and, for $n = 2$, the application of the inner limit, repeated once, yields g_0.

More generally, a repeated application of the inner limit to the outer expansion gives the inner-limit expansion and, by a regrouping of terms, the inner expansion. For instance, for $n = 3$,

$$f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + 2\varepsilon^2 \ln \varepsilon E_2 = 1 + \left(-\frac{1}{x^*} - \varepsilon \ln x^* - \varepsilon \ln \varepsilon + \varepsilon(1 - \gamma) + O(\varepsilon^2) \right)$$

(5.49) $$+ \left(-\frac{\varepsilon \ln x^*}{x^*} - \frac{\varepsilon \ln \varepsilon}{x^*} + \frac{\varepsilon(\gamma - 1)}{x^*} + O(\varepsilon^2) \right)$$

$$+ \left(\frac{2\varepsilon \ln \varepsilon}{x^*} + O(\varepsilon^2 \ln \varepsilon) \right)$$

from which the leading terms of the inner expansion may be computed. (Note that a switchback term of order $\varepsilon^2 \ln \varepsilon$ has been added to the outer expansion in order to match the switchback term of ord $(\varepsilon \ln \varepsilon)$ of the inner expansion.) This implies in particular that each partial outer expansion satisfies the boundary condition at $x = \varepsilon$ to a certain order, a fact which can be verified. We may then ask why one uses the inner expansion at all. One reason is that if one is only interested in calculating conditions near the sphere, then, so to speak, the inner expansion is one step ahead of the outer expansion. A typical quantity to compute would be, say, $\varepsilon(dy/dx)_{x=\varepsilon}$ which, in our physical analogy, may be the heat transfer at the surface of the sphere. Let us take the case $n = 2$. The quantity mentioned may be computed to lowest order from g_0, which is obtained by solving a relatively simple equation and then matching with $f_0 = 1$. Computation of f_1 improves the calculation. (In our calculations we found a better value of α which enabled us to choose $g_1 = 0$.) To the next order we need f_2 to calculate g_2. However, inspection of the calculations shows that we only need the value at zero of f_2 minus its

singular part. This value is the α^2 correction to the temperature at the apparent infinity discussed above. This is exactly the way Kaplun (1957) calculated an improved value of the drag on a circular cylinder in very viscous flow. In this case it would have required extensive numerical calculation to find f_2 as a function of x, but the order unity term of f_2 in the expansion near $x = 0$ could be found by evaluating an integral.

The fact that the outer expansion contains the inner expansion is a coincidence. It will no longer be true if we modify Example II to Example III which will now be studied.

6. Third model equation. As Example III we take a function defined by

(6.1a)
$$\frac{d^2y}{dx^2} + \frac{n-1}{x}\frac{dy}{dx} + \left(\frac{dy}{dx}\right)^2 + y\frac{dy}{dx} = 0,$$

(6.1b)
$$y = 0 \quad \text{at } x = \varepsilon,$$

(6.1c)
$$y = 1 \quad \text{at } x = \infty.$$

This example differs from Example II only by the addition of the term $(dy/dx)^2$ in the left-hand side of the equation.

The approximating equations are

(6.2a)
$$y\frac{dy}{dx} = 0 \quad \text{for } \eta \gg 1,$$

(6.2b)
$$\frac{d^2y}{dx^2} + \frac{n-1}{x}\frac{dy}{dx} + \left(\frac{dy}{dx}\right)^2 = 0 \quad \text{for } \eta \ll 1,$$

and the full equation for ord η = ord 1.

The solution of (6.2b) satisfying the inner boundary condition is

(6.3a)
$$n = 2: \quad g_0 = \ln(1 + B\ln x^*),$$

(6.3b)
$$n = 3: \quad g_0 = \ln\frac{x^* - B}{x^*(1 - B)}.$$

Since the inner equation (6.2b) is nonlinear, the constant of integration B is no longer multiplicative as it was in the corresponding solutions for Example II. This is a significant difference between the two cases.

As in § 5 we show that the leading term of the outer expansion is $f_0 \equiv 1$. The equation for f_1 is the same as in § 5, namely (5.11). The equation for f_2 has an additional forcing term (due to the term $(dy/dx)^2$ in (6.1)) compared to that of § 5, but it is just as easily solved as before.

While there are some great similarities between the two examples, there are some essential differences: (i) The "Stokes equation" (6.2b) is now nonlinear and (ii) is not contained in the "Oseen equation."[22] A consequence of the second difference is that the inner expansion can no longer be retrieved from the outer expansion as was done in (5.49). The important consequence of the first difference will be discussed below.

---

[22] The same mathematical differences exist between the problems of flow at low Reynolds numbers in the compressible case and the incompressible case respectively. (See footnote 16.)

Thus, the leading terms of the outer expansion are

$$f_0 = 1 \quad \text{for all positive integers},$$

(6.4a) $$f_1 = -AE(x) \quad \text{for } n = 2,$$

(6.4b) $$f_1 = -AE_2(x) \quad \text{for } n = 3.$$

For matching we need the expansion of g_0 for large x^*. In the case $n = 2$ we see that $B \ln x^*$ cannot be large since otherwise $\ln \ln x^*$ would be the dominant term. For g_0 to tend to one it is necessary for $B \ln x^*$ to tend to $e - 1$. This gives

(6.5a) $$n = 2: \quad B = \varphi(\varepsilon)(e - 1), \quad \varphi(\varepsilon) = -1/\ln \varepsilon.$$

The domain of overlap is the same as in Example II.
    Similarly,

(6.5b) $$n = 3: \quad B = (e - 1)/e.$$

For $n = 3$, g_0 thus happens to satisfy the condition at infinity; for $n = 2$, it does not.
    In Example II for $n = 2$, the second term of the inner-limit expansion matched with f_0. An important consequence of the nonlinearity of (6.2b) is that this is no longer true for the present example. The inner-limit expansion of $g_0(x^*)$ for $n = 2$ starts

(6.6) $$g_0(x^*) \simeq 0 + \varphi(\varepsilon)(e - 1) \ln x^* + \cdots .$$

We may assume that the terms given explicitly also represent the beginning of the inner-limit expansion of y. We notice now that in the present case *the inner-limit expansion and the outer-limit expansions do not match*; more precisely, no partial sum of the inner-limit expansion of y will match with f_0. The failure of matching is not due to the omission of some third distinguished limit representing some buffer layer. The inner and outer limits are the only distinguished limits. It is plausible a priori that for $n = 2$ the inner limit of y is zero. Hence one may assume the inner-limit expansion to be of the form

(6.7) $$y \simeq 0 + \varphi(\varepsilon)l(x^*) + \cdots ,$$

where $\varphi(\varepsilon) = o(1)$ but is as yet undetermined. The equation for $l(x^*)$ is then

(6.8) $$\frac{d^2l}{dx^{*2}} + \frac{1}{x^*}\frac{dl}{dx^*} = 0.$$

However, unlike (6.2b) this equation is not valid in a domain which overlaps with the outer limit. It is valid in the region where $y = o(1)$, a region which does not overlap with the region where $y \sim 1$ is valid.
    Note that formally a solution of (6.8) matches with $y = 1$, namely,

(6.9) $$\varphi(\varepsilon)l(x^*) = \frac{-1}{\ln \varepsilon} \ln x^*.$$

This is of course the wrong result, since the factor $e - 1$ is missing.[23] Thus, this illustrates the fact that formal success of matching is no guarantee of correctness of the result.

6.1. Change of coordinates for $n = 2$. Bush (1971) has introduced an interesting change of coordinates for $n = 2$ such that the inner expansion actually becomes an inner-limit expansion. He puts

$$(6.10) \qquad\qquad t = \ln (1/x).$$

Equations (6.1) then become ($n = 2$)

$$(6.11a) \qquad\qquad \frac{d^2 Y}{dt^2} + \left(\frac{dY}{dt}\right)^2 - \exp(-t) Y \frac{dY}{dt} = 0,$$

$$(6.11b) \qquad\qquad Y = \begin{cases} 0 & \text{at } t = \ln(1/\varepsilon), \\ 1 & \text{at } t = -\infty, \end{cases}$$

where $y(x, \varepsilon) = Y(t, \varepsilon)$.

The transformation is important only for the inner expansion, and we shall restrict our discussion accordingly. The inner layer occurs at t large, and Bush introduces the inner variable

$$(6.12) \qquad\qquad t^* = \varphi(\varepsilon) t.$$

The last term in (6.11a) will be exponentially small and hence transcendentally small in the inner region (which is especially easy to see in the present coordinates, cf. statement following (5.34)). We put

$$(6.13) \qquad\qquad g_j(x^*, \varepsilon) = G_j(t^*)$$

(as will be seen, the G_j will not depend on ε explicitly). The inner expansion is then

$$(6.14) \qquad\qquad Y \simeq G_0 + \varphi G_1 + \cdots.$$

One finds

$$(6.15a) \qquad\qquad G_0 = \ln\left[1 + (e - 1)(1 - t^*)\right],$$

$$(6.15b) \qquad\qquad G_1 = \gamma(e - 1)(1 - t^*)[1 + (e - 1)(1 - t^*)]^{-1}.$$

(The computation of G_1 necessitates matching with the third outer term $f_2(x)$ which is computed in the same way as the corresponding f_2 of Example II in § 5. Details are given in Bush (1971).)

Thus G_0 and G_1 do not contain ε explicitly and each term in the inner expansion is now identical with the corresponding term in the inner-limit expansion. This fact does not change any of the heuristic ideas of matching. However, it has certain advantages for the actual construction of the expansion. Bush's procedure is as follows. One first finds g_0 as shown earlier. Guided by the form of g_0 one finds the variables t and t^* such that the explicit dependence of ε will

---

[23] It has been proved rigorously in Cohen and Lagerstrom (to be published) that g_0 as given by (6.3a) is indeed correct.

disappear in $G_0(t^*)$. The equation for G_1 is then expected to be somewhat simpler than that for g_1. In the present case we get

(6.16)
$$\frac{d^2 G_1}{dt^{*2}} + 2\frac{dG_0}{dt^*}\frac{dG_1}{dt^*} = 0.$$

Defining g_1 as the function of x and ε which corresponds to $G_1(t^*)$ we see from (6.16) that it obeys the equation

(6.17a)
$$\frac{d^2 g_1}{dx^{*2}} + \frac{1}{x^*}\frac{dg_1}{dx^*} + 2\frac{dg_1}{dx^*}\frac{dg_0}{dx^*} = 0$$

with

(6.17b)
$$\frac{dg_0}{dx^*} = \frac{\varphi(\varepsilon)(e-1)}{[1 + (e-1)\varphi(\varepsilon)\ln x^*]x^*}.$$

The solution of (6.16), satisfying appropriate matching and boundary conditions, is

(6.18)
$$G_1 = \frac{\gamma(e-1)(1-t^*)}{1 + (e-1)(1-t^*)}.$$

Rewriting this in x^* one finds the solution of (6.17) to be

(6.19)
$$g_1 = \frac{\gamma(e-1)\varphi\ln x^*}{1 + (e-1)\varphi\ln x^*}.$$

For large values of x^*, one finds, using the outer variable, that

(6.20)
$$g_1 = \gamma\frac{e-1}{e} + \varphi\gamma\frac{e-1}{e^2}\ln x + O(\varphi^2).$$

From the point of view of the present paper, the interest in Bush's transformation is not the simplification of the calculations, but rather the following two points: First, it is not necessary to form the inner variable by dividing x by a function of ε; one may instead divide some function of x by a function of ε. Instead of a function of x one may even use a more general coordinate change as described by (3.7). This is a rather obvious statement, but it is also a nonconstructive generality. The interest in Bush's analysis lies in the fact that a *specific* coordinate change has actually been found systematically which simplifies the analytical work, and the question is raised whether this may be done in a systematic way for a large class of problems. Secondly, and this is of special interest in the context of the present paper, the following problem is raised by the coordinate change: (6.17a) is obtained by applying limit processes to the proper equations, as shown in detail in § 5, provided the explicit formula for the derivative of g_0 is not used. However, if the explicit expression (6.17b) is used, then the third term will disappear in the limit and the resulting equation is

(6.21)
$$\frac{d^2 \tilde{g}_1}{dx^{*2}} + \frac{1}{x^*}\frac{d\tilde{g}_1}{dx^*} = 0.$$

The solution[24] is

$$(6.22) \qquad \tilde{g}_1 = \varphi\gamma\frac{e-1}{e}\ln x^* = \gamma\frac{e-1}{e} + \varphi\gamma\frac{e-1}{e}\ln x.$$

The two results are significantly different. As was the case for Example II, in a physical problem $(dy/dx^*)_{x^*=1}$ may be important. The contribution of φg_1 to this quantity is $\varphi^2\gamma(e-1)$ whereas $\tilde{g}_1$ would give $\varphi^2\gamma(e-1)/e$. Thus the difference between the two values cannot be corrected by higher order terms. However, g_1 is the correct answer; the reason for this will now be discussed.

6.2. Integrated effects. To examine the situation more closely we write (6.17a) as

$$(6.23) \qquad \frac{d}{dx^*}\left[\ln\left(x^*\frac{dg_1}{dx^*}\right)\right] = -2\frac{dg_0}{dx^*}.$$

Inserting (6.17b) we find that the right-hand side is of order φ. However, integrating transforms it into a term of order unity. Thus the integrated effect is of order unity and the function g_1 rather than $\tilde{g}_1$ gives the correct answer.

In § 5, after (5.26) we discussed the integrated effects of a term of the form (function of ε) · (function of x). We observed that the effect of such a term could be larger than anticipated if a matching condition rather than a boundary condition were imposed. In that case there was no doubt about which equation to use for g_1 (or rather for g_0). The only question was which solution to use; the correct solution was not the one expected from a limit-process expansion or from a superficial order estimate. In the present case the main question is which equation to use. We actually have to modify our notion of formal limit of an equation. Since according to (5.32) the η-limit of $\varphi\ln x^*$ is $1 - a$ for $\eta = \varepsilon^a$, a formal application of limits to (6.23) would yield infinitely many equations. It is therefore natural to extend the discussion after (5.43) from solutions to equations, i.e., to include all limits we have to regard $\varphi\ln x^*$ as formally of order unity. We also note that in the present case g_1 and g_0 are of the same order.

From Examples II and III we learn that the heuristic principle (4.32) is a very useful starting point in constructing solutions but that it needs to be modified by the consideration of integrated effects. Furthermore we have just seen that our original ideas about limits of equations have to be modified. The statement that integrated effects should be considered correctly is, of course, a tautology. However, our aim here has been to give specific examples of how large integrated effects occur and how one can spot them. We believe that the examples given are typical of a large class of problems. The easiest one to spot is the switchback effect which occurred in Example II for $n = 3$. Here impossibility of matching warned us that the assumed form of the expansion was wrong. The difficulties illustrated by Example III are more subtle. It shows that formal success of matching is no guarantee that the solutions actually do match. This was first illustrated by the incorrect formula for g_0 given by (6.9). However, in

---

[24] Both g_1 and $\tilde{g}_1$ match with the "temperature" correction at the apparent infinity as given by the constant term $\gamma(e-1)/e$ of f_2. The discussion after (5.43) applies equally well to the present case.

this case our heuristic principle gave the correct result and there are no large integrated effects. The difficulties in finding g_1 are greater. Suppose that (6.17a) had been given as the original equation with dg_0/dx^* written out explicitly and that the matching and boundary conditions are as above. If one wanted to find a first approximation for ε small, how could one see that g_1 rather than $\tilde{g}_1$ is the correct result? The only answer seems to be that once one poses the question one can find the right result. However, if one is not forewarned of the possibility of large integrated effects and blindly neglects terms of order φ, then the wrong answer will result.

6.3. Matching reconsidered. As discussed at the end of § 4, the older technique of interchanging limit expansions may be justified in certain simple cases and where it was shown to be in essence identical with the use of the intermediate limit. In simple cases one may even formulate a precise rule such as (4.61). Van Dyke (1964, p. 90) has proposed a generalization of this rule as a general matching principle. In a careful study Fraenkel (1969) has shown conditions under which this rule holds and also given some counterexamples.

Although the present paper has a purpose quite different from that of Fraenkel's, some comments will be made. A basic difficulty with *any* general matching principle based on interchange of inner and outer limit processes is that it presupposes that the inner and outer expansions actually are obtained by limit processes. The idea of characterizing expansions by their domain of validity is not only more basic for understanding matching but often essential in constructing these expansions. An obvious example of this is given in § 6. Less obvious but equally important is Example II for $n = 3$. As discussed after (5.46) it is really the function $-\varepsilon \ln \varepsilon g_0(x^*) + \varepsilon g_1(x^*)$ which is the second term of the inner expansion. One may of course try to formulate rules about applying the inner-limit expansion to the outer expansion (in the sense used here) and comparing it with the outer-limit expansion of the inner expansion. Any such rule, which gives the number of terms to be used in the inner and outer expansions, must specify the system for counting terms.

Another important point, which has been made earlier, is that whatever techniques are used, matching is always based on overlap. It obviously does not make sense to compare two approximations which do not have a common domain of validity. Furthermore, the technique always consists of comparing the outer expansion for small x with the inner expansion for large x^* (assuming of course that x and x^* are the two principal variables). The present paper shows by several examples how this can be carried out very simply by considering the order of the terms in the difference between a partial outer and a partial inner expansion.

7. Examples from fluid dynamics. In previous sections some connected problems of singular perturbations were discussed with the aid of relatively simple model equations. Although these equations were originally introduced to illustrate the use of certain ideas and techniques of singular perturbations used in mathematical fluid dynamics, they may be discussed without reference to fluid dynamics. In this section we shall show how a variety of special problems may arise in singular perturbations using examples taken directly from fluid dynamics. In many of the cases it seems difficult to invent simple model equations

without losing the essence of the problem. Further examples are found in Van Dyke (1964).

7.1. The Navier–Stokes equations. We shall be concerned mainly with two-dimensional stationary incompressible viscous flow (see Lagerstrom (1964)). From

(7.1a) $v =$ kinematic viscosity of the fluid,

(7.1b) $u_0 =$ reference velocity,

(7.1c) $p_0 =$ reference pressure,

one may form the reference length of the fluid

(7.2) $\lambda = v/u_0$.

We shall let $\mathbf{x} = (x_1, x_2)$ denote the rectangular length coordinate nondimensionalized by λ. The velocity nondimensionalized by u_0 will be denoted by $\mathbf{q} = (q_1, q_2)$. The pressure, nondimensionalized by p_0, will be denoted by p.

The Navier–Stokes equations in nondimensional form are then

(7.3a) $$\sum_{i=1}^{2} q_{i,i} = 0,$$

(7.3b) $$\sum_{i=1}^{2} (q_i q_{j,i} + p_{,j} - q_{j,ii}) = 0, \qquad\qquad j = 1, 2.$$

Here $f_{,i} = \partial f/\partial x_i$.

As boundary conditions we require that

(7.4) $\mathbf{q} = 0$ on solid surface.

For flow of an infinite medium past a solid we require

(7.5) $\mathbf{q} = 1, \quad p = 1$ at infinity.

This means, of course, that we have used the value at infinity as reference value in (7.1b) and (7.1c).

The boundary condition (7.4) refers to a solid surface. If this surface is a characteristic length L, we may form a nondimensional parameter, called the Reynolds number and denoted here by R,

(7.6) $R = L/\lambda$.

As discussed in § 2, if in a problem the two lengths are of different order of magnitude, we may use perturbation techniques but must expect that the problems may be singular.

7.2. Flow at low Reynolds numbers. Consider viscous flow past an object whose typical dimension L is much smaller than the viscous length λ, i.e.,

(7.7) $L \ll \lambda \quad \text{or} \quad R \ll 1$.

As inner coordinate we use

(7.8) $\mathbf{x}^* = \mathbf{x}_{\text{dim}}/L = \mathbf{x}/R$.

The basic ideas on how to construct an asymptotic expansion for small R are due to Kaplun. They will not be discussed here (see Van Dyke (1964) and references given there). The essentials of the ideas involved are shown by our second model equation. (The relation of the second and third model equations to fluid dynamics is discussed in footnote 16.)

7.3. Flow at high Reynolds numbers. Written in x_j^*-coordinates, (7.3) becomes

$$\text{(7.9a)} \qquad \sum \frac{\partial q_i}{\partial x_i^*} = 0,$$

$$\text{(7.9b)} \qquad \sum_i q_i \frac{\partial q_j}{\partial x_i^*} + p_{,j} = \frac{1}{R} \sum_i \frac{\partial^2 q_j}{\partial x_i^{*2}}.$$

For large values of R we put the right-hand side of (7.9b) equal to zero. The resulting equations are called the *Euler equations*. However, due to the lowering of the order of the equations we have to weaken (7.4) to

$$\text{(7.10)} \qquad \text{normal component of } \mathbf{q} = 0 \quad \text{on solid surface.}$$

The viscous force exerted by the fluid on the body, the so-called skin friction, cannot be computed from a solution of (7.10). Prandtl (1905) pointed out that the tangential velocity changes very rapidly from its true value ($= 0$) to the value given by the solution of (7.10). This change takes place in a layer, called the boundary layer, of thickness ε, where

$$\text{(7.11)} \qquad \varepsilon = 1/\sqrt{R}.$$

To study this layer, Prandtl replaced the x_i^* by coordinates ξ and η, where $\eta = 0$ on the surface of the body. If then η is replaced by

$$\text{(7.12)} \qquad \bar{\eta} = \eta/\varepsilon$$

and, using our terminology, the formal limit for $\varepsilon \downarrow 0$ is applied to the resulting equations, one obtains Prandtl's boundary-layer equations. If ξ and η are coordinates tangential and normal to the surface of the solid respectively, and $d\xi$ and $d\eta$ give the correct length element (relative to the x_i^*) at the surface, and if u and v are the velocity components relative to ξ and η, then the Prandtl equations are

$$\text{(7.13a)} \qquad \frac{\partial u}{\partial \xi} + \frac{\partial v}{\partial \eta} = 0,$$

$$\text{(7.13b)} \qquad u\frac{\partial u}{\partial \xi} + v\frac{\partial u}{\partial \eta} + \frac{\partial p}{\partial \xi} = \frac{1}{R}\frac{\partial^2 u}{\partial \eta^2},$$

$$\text{(7.13c)} \qquad \partial p/\partial \eta = 0.$$

As boundary conditions Prandtl proposed the obvious one, namely (7.4), and the more surprising one:[25]

---

[25] This corresponds to (4.9c). It was precisely to illustrate some basic aspects of Prandtl's perturbation method that Friedrichs (1953) introduced a model equation similar to our Example I of §4.

(7.14)

The values of p and u of the boundary-layer solution at $\bar{\eta} = \infty$ equal the corresponding values (denoted by p_e and u_e) of the Euler solution at $\eta = 0$.

We note that from the Euler equations it follows that

$$(7.15a) \qquad \frac{d}{d\xi}\left(\frac{u_e^2}{2} + p_e\right) = 0,$$

and, since a physically irrelevant constant may be added to the pressure,

$$(7.15b) \qquad u_e^2/2 + p_e = 0.$$

For a detailed discussion of boundary-layer theory and higher order approximations from the point of view of asymptotic theory, see, for instance, Lagerstrom (1964) and Van Dyke (1964). Some specialized problems will be discussed below.

7.4. Role of subcharacteristics. In a singular perturbation problem involving partial differential equations the characteristic lines (or surfaces) of the outer equations are called subcharacteristics. In our case the outer equations are the Euler equations whose only real characteristics are the streamlines, i.e., the lines everywhere tangent to the velocity field. Because of (7.10) the boundary is a streamline, and on this streamline there is a discontinuity between the prescribed values of the tangential velocity and the ones given by the Euler solution. Latta (1951) showed for various partial differential equations, that the thickness of the boundary layer depends on whether the discontinuity occurs on a subcharacteristic or not. The application of this principle to fluid dynamics is discussed in detail by Lagerstrom (1964). A simple illustration is the following. If the first two terms on the left-hand side of (7.9b) (the so-called transport terms) are replaced by $\partial q_j/\partial x_1^*$, the so-called Oseen equations result. The characteristic lines are now the lines $x_2^* = \text{const}$. The thickness of the boundary layer is $R^{-1/2}$ when the boundary is parallel to the x_1^*-axis; it is R^{-1} when the boundary has positive slope; and when the boundary has negative slope no boundary layer exists (loc. cit.). A general discussion of the role of subcharacteristics is given by Cole (1968, Chap. 4).

7.5. Nonuniqueness of the Euler solution. Actually the asymptotic theory for flow at high Reynolds numbers leads to some very difficult and as yet unsolved problems. A crucial difficulty is that the solution of the Euler equations with the boundary conditions given above is not unique, and it is often difficult to select the relevant solution, i.e., the one which is the limit of the Navier–Stokes equations as $R \to \infty$.

In some cases a study of the vorticity is helpful. We define

$$(7.16) \qquad \text{vorticity} = \omega = \partial q_1/\partial y - \partial q_2/\partial x.$$

It is then easily seen that vorticity obeys the equation

$$(7.17) \qquad q_1\frac{\partial\omega}{\partial x_1^*} + q_2\frac{\partial\omega}{\partial x_2^*} = \frac{1}{R}\left(\frac{\partial^2\omega}{\partial x_1^*} + \frac{\partial^2\omega}{\partial y_2^*}\right).$$

A curve everywhere tangent to $\mathbf{q}$ is called a streamline. Equation (7.17) may then be interpreted as follows: Vorticity is transported by the streamlines and at

the same time it diffuses like heat with coefficient $1/R$. In the limit of infinite R vorticity is constant on each streamline. For flow past an object the boundary conditions prescribe $q_1 = 1$ and $q_2 = 0$, and hence $\omega = 0$ at $x_1 = -\infty$. If every streamline comes from upstream infinity, then in the limit vorticity is zero everywhere. This assumption reduces the Euler equations to the Laplace equation which has a unique solution. For semi-infinite bodies whose width increases monotonely downstream it seems plausible that every streamline comes from upstream infinity. The simplest example of such a body is the parabola with axis along the positive x_1-axis. For this case Van Dyke (1964a) has computed the two leading terms of the inner and outer expansions respectively.

However, for the case of flow past a finite body the assumption that the limiting solution obeys the Laplace equation leads to contradictions with boundary-layer theory (see Lagerstrom (1964)). The limiting flow must then contain streamlines with nonzero vorticity; these lines must either be closed or come from and return to downstream infinity ($x_1 = \infty$). A limiting case is the so-called Kirchhoff flow where two streamlines, originating at the solid, separate a region of potential flow from a region of zero velocity; these streamlines carry infinite vorticity. However, so far no valid arguments have been established to characterize the limiting flow in such a way as to make the solution of the Euler equations unique. The difficulties are connected with the very complex problem of flow separation and nonexistence of solutions of the Prandtl equations, subjects which cannot be discussed here.

We notice that unlike examples discussed earlier, the leading term of the outer expansion cannot be determined a priori, in other words, we do not know what is being perturbed. A solution would involve a simultaneous consideration of the leading term of both the inner and the outer expansion (and possibly other expansions). A similar situation exists for the problem discussed below which, however, is much simpler and which can be solved for special boundary conditions.

7.6. Limiting flow inside a finite domain. We consider a finite domain D with boundary B. In the boundary conditions for (7.3) we replace (7.4) by

(7.18a)   normal velocity component $= 0$,

(7.18b)   tangential velocity component $=$ given function u_b.

For the Euler flow in the same domain we use (7.18a) and give up (7.18b). Obviously the solution is not unique. If $(\mathbf{q}, p)$ is one solution, then $(k\mathbf{q}, k^2 p)$ with $k = $ const. is another solution and many other drastically different solutions exist. However, the limiting flow is characterized by the following property (Prandtl (1905)):

(7.19)   The vorticity is constant in each vortex.

To define a vortex we say that two streamlines belong to the same vortex if and only if one encloses the other and if the region between the streamlines consists entirely of such "nested" streamlines. Thus the flow depicted in Fig. 7.1 consists of two vortices. The bounding streamlines are ABC and ACD respectively. The streamline $ABCD$ does not belong to either vortex.
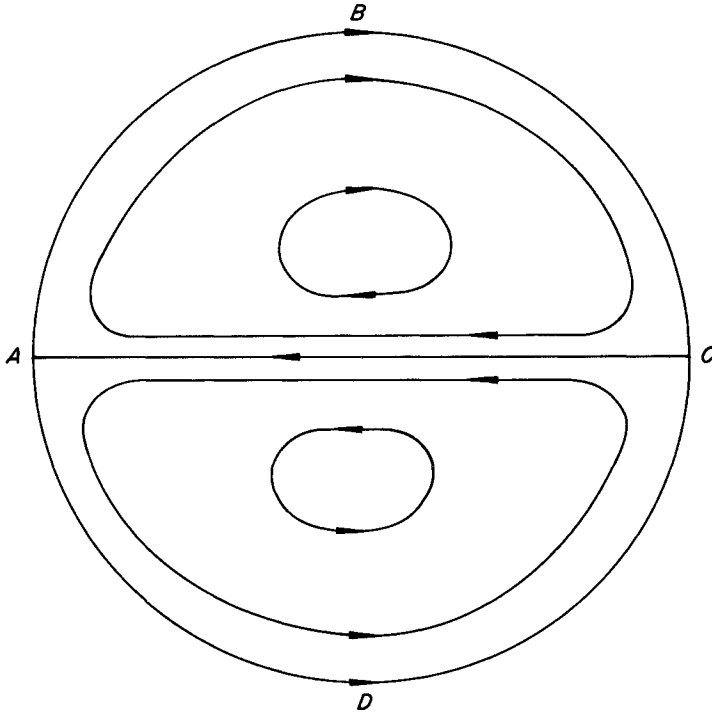
FIG. 7.1. *Vortices*

The proof of (7.19) depends on the assumption that in a finite domain the streamlines are closed (which follows if one assumes that there are no mass-sources or mass-sinks in the domain).

We may write the two equations of (7.9b) as

$$(7.20a) \qquad \frac{-\partial h}{\partial x_2^*} + q_1\omega = \frac{1}{R}\frac{\partial \omega}{\partial x_1^*},$$

$$(7.20b) \qquad \frac{\partial h}{\partial x_1^*} + q_2\omega = \frac{1}{R}\frac{\partial \omega}{\partial x_2^*},$$

where

$$(7.20c) \qquad h = \tfrac{1}{2}(q_1^2 + q_2^2) + p.$$

Consider a closed streamline S with normal $d\mathbf{n} = (-dx_2^*, dx_1^*)$. Multiply (7.20a), (7.20b) by dx_2^* and dx_1^*, respectively, and integrate over the streamline. The integral of the first term gives the increase in h going around a closed curved and is zero since h is one-valued. The second term gives zero since $\mathbf{q} \cdot d\mathbf{n}$ is zero. Thus we obtain

$$(7.21) \qquad \oint_S \frac{\partial \omega}{\partial n}\, ds = 0,$$

where ds is the length element along the curve S. This result is valid for arbitrary R.

We now introduce the streamfunction ψ by

$$(7.22) \qquad q_1 = \partial\psi/\partial x_2^*, \qquad q_2 = -\partial\psi/\partial x_1^*,$$

a definition which is possible because of (7.9a). Obviously the value of ψ does not change along a streamline. In the limit $R \to \infty$, the value of ω is also constant along a streamline so that $\omega = \omega(\psi)$ and furthermore $d\omega/d\psi$ is constant along each streamline. We then rewrite (7.21) as

$$(7.23) \qquad \left(\frac{\partial\omega}{\partial\psi}\right)_s \oint \frac{\partial\psi}{\partial n}\, ds = 0.$$

Since, as follows from (7.22), $\partial\psi/\partial n$ is the tangential velocity component, the value of the integral is nonzero. The conclusion presupposes that S is part of a vortex (it would be invalid for the streamline $ABCD$ in Fig. 7.1). Hence

$$(7.24) \qquad \partial\omega/\partial\psi = 0 \quad \text{for } R = \infty.$$

This concludes the proof of (7.19).

Returning to the perturbation problem we consider first the flow inside a circle of radius unity in x_j^*-coordinates. The prescribed tangential velocity component u_b is now a function of the angle θ. We shall see if (7.19) will help in finding the solution. The obvious difficulty is that we do not know how many vortices there are. *Assume* tentatively that only one vortex occurs. It can easily be shown that the only solution of the Euler equations with constant vorticity is when the fluid rotates like a solid, i.e., when the r-component of the velocity is zero and the θ-component, denoted by u, is proportional to r. This still leaves a one-parameter family of solutions: If (u, p) is a solution, then so is $(ku, k^2 p)$. The flow field is uniquely determined if u_e is given, where

$$(7.25) \qquad u_e = u \quad \text{at } r = 1.$$

Thus we want to relate the unknown u_e to the known $u_b(\theta)$ through a boundary-layer solution. We note that the *Euler flow* at $r = 1$ is u_e whereas the actual velocity field at $r = 1$ is u_b. We introduce the boundary-layer coordinates (cf. (7.12))

$$(7.26) \qquad \xi = \theta, \quad \eta = 1 - r, \quad \bar{\eta} = \eta\sqrt{R}.$$

As boundary conditions for the boundary-layer flow we then find

$$(7.27a) \qquad u = u_b \quad \text{at } \bar{\eta} = 0,$$

$$(7.27b) \qquad u = u_e \quad \text{at } \bar{\eta} = \infty.$$

The criterion for finding u_e is that the Navier–Stokes solution is periodic in θ and that hence the boundary-layer solution must be periodic in $\xi(= \theta)$. Our problem is then to translate this criterion into an a priori estimate of u_e.

For the circle one proceeds as follows (Feynman and Lagerstrom (1956)).

We make the coordinate change from (ξ, η) to (ξ, ψ) (known to aerodynamicists as von Mises coordinates). The boundary-layer equation (7.13b) then becomes

$$(7.28) \qquad u\frac{\partial u}{\partial\xi} + \frac{\partial p}{\partial\xi} = u\frac{\partial^2(u^2/2)}{\partial\psi^2}.$$

This equation is valid when p_e is an arbitrary function of ξ. In the present case p_e is constant and (7.28) reduces to

$$2\frac{\partial u}{\partial \xi} = \frac{\partial^2 u^2}{\partial \psi^2}.$$

(7.29)

We define the average of a function $f(\xi, \psi)$ by

$$\bar{f}(\psi) = \frac{1}{2\pi} \int_0^{2\pi} f(\xi, \psi) \, d\xi, \qquad \psi = \text{const}.$$

We can now use periodicity: $u(\xi, \psi) = u(\xi + 2\pi, \psi)$ implies that the average of the left-hand side (of (7.29)) is zero. Then[26]

$$d^2 \bar{u}^2 / d\psi^2 = 0.$$

(7.30)

Thus the derivative of the mean of u^2 does not vary from streamline to streamline. Evaluation at $\bar{\eta} = \infty$ shows that this derivative actually is zero on every streamline. Thus $\bar{u}^2$ is the same on all streamlines. Evaluation at $\bar{\eta} = \infty$ and $\bar{\eta} = 0$ gives, finally,

$$u_e^2 = \overline{u_b^2}.$$

(7.31)

Hence we have found an a priori relation between the given function $u_b(\theta)$ and the unknown constant u_e.

Equation (7.31) does not give the sign of u_e. There is, however, a much more serious difficulty than that. Equation (7.31) is a necessary but not a sufficient condition for the existence of a boundary-layer solution. Consider the Euler flow as in Fig. 7.1 and assume that $\omega = +1$ in the upper half of the circle and $\omega = -1$ in the lower half of the circle. If one prescribes u_b to be equal to the values taken by the Euler flow on the boundary or deviating very slightly therefrom while still keeping the basic symmetry around AC, the limiting flow must be of the form shown in Fig. 7.1 with possibly a slightly different value of $|\omega|$. Thus computation of a u_e from (7.31) is meaningless in this case.

The vortex picture shown in Fig. 7.1 is still relatively simple; there may be a large number of vortices in an asymmetrical arrangement.

At present the limiting vortex structure cannot be determined a priori. Thus, even for the flow inside a circle the problem is not solved. However, if the resulting flow has only one vortex, the solution may be obtained with the aid of (7.31). The natural way to test whether there is one vortex is the following:

One computes $|u_e|$ by (7.31) and gives u_e the same sign as the average of u_b and studies the resulting boundary-layer equations. The breakdown of the boundary-layer solution indicates that there is more than one vortex. The theory of the boundary-layer equations furnishes approximate criteria for the non-existence of a solution. Clearly, if u_b differs only slightly from the computed u_e, there will be only one vortex.

If the boundary is not a circle, the equivalent of (7.31) is not known. However, assume that the domain is simply connected. Let the given boundary condition u_b have the form

$$u_b = f(\xi) + \varepsilon g(\xi),$$

(7.32)

---

[26] The argument used amounts to the consideration of momentum balance over one period.

where $f(\xi)$ is the value at the boundary of a special one-vortex Euler solution (which can be found by solving Poisson's equation). We assume ε to be small. Then the actual Euler flow corresponding to the boundary conditions (7.32) differs from the special one by a multiplicative constant k. Thus

$$(7.33) \qquad\qquad u_e = k \cdot f.$$

For $\varepsilon = 0$, $k = 1$. Hence

$$(7.34) \qquad\qquad k = 1 + \varepsilon k_1 + \cdots .$$

The periodicity requirement then gives (Feynman and Lagerstrom (1956))

$$(7.35) \qquad\qquad k_1 = \overline{f^2 g}/\overline{f^3}.$$

This result reduces to (7.31) when the boundary is a circle. It may be proved using von Mises coordinates. The proof is similar to that of (7.31).

7.7. Optimal coordinates. By some ingenious geometrical reasoning Kaplun (1954) showed that if one finds the first two terms of the outer expansion and the first term of the inner expansion for flow at high Reynolds numbers in two dimensions, then one may introduce a new system of coordinates, called optimal coordinates, such that the leading term of the inner expansion with respect to the new system contains the two leading terms of the outer expansion. This work has been extended to higher order by Legner (1971). Thus, up to the order considered, the two matching expansions have been replaced by one regular expansion.

7.8. Coordinate-type expansions, artificial parameters. In § 3 it was mentioned that the distinction between coordinates and parameters is not intrinsic but depends on the problem studied and the method of solution. In the asymptotic expansions studied here we have sought expansions which are uniformly valid in a closed domain of coordinate-space as the parameter ε tends to a limiting value. The quantity ε was also a parameter in the sense that derivatives with respect to ε never appeared in the differential equations determining the function. There are however cases where a variable is physically a coordinate and occurs like a coordinate in the differential equation, but in the asymptotic solution we treat it as a parameter. We may for instance consider an expansion of $f(x_1, x_2, \varepsilon)$ uniformly valid in x_2 as x_1 tends to infinity and for a fixed value of ε. Such expansions are called coordinate-type expansions as opposed to the parameter-type expansions studied earlier; see Lagerstrom and Cole (1955) and Chang (1961).

A typical problem of coordinate-type expansions occurs in the theory of viscous compressible flow past a finite body, the problem being to find the flow field at a large distance from the body. The two-dimensional case of this problem has been studied thoroughly by Chang (1961) who also discusses the general methodology for coordinate-type expansions. The problem is singular; in the narrow wake region downstream of the body called the wake, a different scale in x_2 is used than outside the wake. The solutions for different regions are matched as in a parameter-type expansion; even switchback occurs. The quantity corresponding to the small parameter ε is x_1^{-1}. One may achieve the same formal result by introducing an artificial dimensional parameter in the form of an artificial

length. The small parameter is then the body dimension divided by the artificial length. Of course, the artificial length must disappear from the final answer (see Chang (1961)).

It must be emphasized that for fixed Reynolds number the flow at large distances cannot be matched with the flow near the body. Since, however, the flow at large distances is determined by the boundary conditions on the body, the asymptotic expansion for large distances will then contain undetermined constants. Assume one also finds a solution near the body, say a Taylor expansion in suitable coordinates. This expansion will also contain undetermined constants. However, the undetermined constants of the two expansions cannot be related by matching, i.e., by comparison of partial sums. The problem is that of analytic continuation rather than matching; only the entire Taylor expansion near the body will determine the asymptotic expansion at large distances. One may, however, use integral conservation laws as a kind of analytic continuation in the mean.[27] For instance, since divergence of the velocity field is zero (conservation of mass) and since the flow through the surface of the body is zero it follows that the integral of $\mathbf{q} \cdot d\mathbf{n}$ over any closed curve is zero. This determines one of the constants of the asymptotic expansion at infinity. Similarly, other constants in this expansion may be related to the total force and the torque exerted by the fluid on the body.

Another problem which has only one length is nonviscous flow past a parabola, or paraboloid. The problem has only one independent real length, say the nose radius of curvature, so that one cannot talk about a slender parabola. However, far downstream the parabola looks slender; i.e., the radius divided by the distance to the nose is small. One may formalize this by introducing a large artificial length, and use singular perturbation techniques. The solution will not be valid near the nose and no matching is possible. However, the law of conservation of mass can be used to great advantage here. Details are given in Van Dyke (1958).

We note that in a given problem a physical parameter may be mathematically artificial depending on the boundary conditions. For instance, in viscous flow past a semi-infinite plate (or any conical body, i.e., body with no length parameter) the viscosity is artificial in the sense that it must be eliminated. From dimensional analysis we know that v must appear only in the combinations (cf. (7.2)) $x_j = $ (dimensional coordinate)$\cdot u_0/v$. Thus if we use these coordinates, v will disappear both from the equations (7.3) and the boundary conditions. One may, of course, expand in small values of v but this is in reality an expansion for large x. The expansion will not be uniformly valid at the nose of the flat plate. More precisely, it is easily shown that if we have any approximation which is uniformly valid to order unity everywhere and where v occurs only in the combination described in the definition of x_j above, then the approximation is in reality the exact solution; see Lagerstrom and Cole (1955, pp. 869–870) and Chang (1961, pp. 821–822).

---

[27] This leads to the interesting question of finding all independent conservation laws for a system of equations. This problem is discussed in Casten (1970).

Appendix. The equation

(A.1a)
$$\frac{d^2f}{dx^2} + \frac{m}{x}\frac{df}{dx} + \frac{df}{dx} = 0$$

with boundary condition

(A.1b)
$$f(\infty) = 0$$

determines f to be a constant times the function

(A.2)
$$E_m(x) = \int_x^\infty e^{-t}t^{-m}\,dt.$$

The corresponding nonhomogeneous equations

(A.3)
$$\frac{d^2f}{dx^2} + \frac{m}{x}\frac{df}{dx} + \frac{df}{dx} + E_m\frac{dE_m}{dx} = 0$$

and

(A.4)
$$\frac{d^2f}{dx^2} + \frac{m}{x}\frac{df}{dx} + \frac{df}{dx} + \left(\frac{dE_m}{dx}\right)^2 = 0$$

have the particular solutions, with $f(\infty) = 0$,

(A.5)
$$f = F_m(x) = \frac{dE_{m-1}(x)}{dx}E_m(x) + 2^{2m-1}E_{2m-1}(2x)$$

and

$$f = G_m(x) = -\tfrac{1}{2}(E_m(x))^2,$$

respectively.

Asymptotic expansions valid for x small are

(A.6a) $\quad E_1(x) = -\ln x - \gamma + x + O(x^2),$

(A.6b) $\quad E_2(x) = \dfrac{e^{-x}}{x} - E_1(x) = \dfrac{1}{x} + \ln x + (\gamma - 1) - \dfrac{x}{2} + \dfrac{x^2}{12} + O(x^3),$

(A.6c) $\quad E_3(x) = \dfrac{e^{-x}}{2x^2} - \dfrac{e^{-x}}{2x} + \dfrac{1}{2}E(x) = \dfrac{1}{2x^2} - \dfrac{1}{x} - \dfrac{1}{2}\ln x + \dfrac{3-2\gamma}{4} + \dfrac{x}{6} + O(x^2),$

(A.7a) $\quad F_1(x) = -\ln x - 2\ln 2 - \gamma - x\ln x + (3-\gamma)x + O(x^2\ln x),$

(A.7b) $\quad F_2(x) = -\dfrac{\ln x}{x} - \dfrac{\gamma+2}{x} - 3\ln x - 3\gamma + 5 - 4\ln 2 + O(x\ln x),$

(A.8a) $\quad G_1 = -\dfrac{1}{2}\ln^2 x - \gamma\ln x - \dfrac{\gamma^2}{2} + x\ln x + \gamma x + O(x^2),$

$$G_2 = -\frac{1}{2x^2} - \frac{1}{x}\ln x - \frac{\gamma-1}{x} - \frac{\ln^2 x}{2}$$

(A.8b)
$$- (\gamma - 1)\ln x\frac{-\gamma^2 + 2\gamma}{2} + \frac{x\ln x}{2} + \frac{\gamma-1}{2}x + O(x^2).$$

REFERENCES

[1] WM. B. BUSH (1971), *On the Lagerstrom mathematical model for viscous flow at low Reynolds number*, this Review, to appear.

[2] R. G. CASTEN (1970), *Methods for deriving conservation laws*, Doctoral thesis, California Institute of Technology, Pasadena.

[3] I. E. CHANG (1961), *Navier–Stokes solutions at large distances from a finite body*, J. Math. Mech., 10, pp. 811–876.

[4] D. S. COHEN AND P. A. LAGERSTROM, *Asymptotic theory of a model equation*, to be published.

[5] J. D. COLE (1968), *Perturbation Methods in Applied Mathematics*, Blaisdell, Waltham, Mass.

[6] A. ERDÉLYI (1956), *Asymptotic Expansions*, Dover, New York.

[7] R. P. FEYNMAN AND P. A. LAGERSTROM (1956), *Remarks on high Reynolds number flows in finite domains*, Proc. IX International Congress on Applied Mechanics, vol. 3 (Brussels, 1956).

[8] L. E. FRAENKEL (1969), *On the method of matched asymptotic expansions*, Proc. Cambridge Philos. Soc., Part I, 65, pp. 209–231. Part II, 65, pp. 233–261; Part III, 65, pp. 263–283.

[9] K. O. FRIEDRICHS (1953), *Special Topics in Fluid Dynamics*, New York University Press, New York; see also VON MISES AND FRIEDRICHS (1953), *Lecture Notes on Fluid Dynamics*.

[10] S. GOLDSTEIN (1960), *Lectures on Fluid Mechanics*, Interscience, New York.

[11] SAUL KAPLUN (1954), *The role of coordinate systems in boundary-layer theory*, Z. Angew. Math. Phys., 5, pp. 111–135.

[12] ——— (1957), *Low Reynolds number flow past a circular cylinder*, J. Math. Mech., 6, pp. 595–603.

[13] ——— (1967), *Fluid Mechanics and Singular Perturbations*, P. A. Lagerstrom, L. N. Howard and C. S. Liu, eds., Academic Press, New York.

[14] SAUL KAPLUN AND P. A. LAGERSTROM (1957), *Asymptotic expansions of Navier–Stokes solutions for small Reynolds numbers*, J. Math. Mech., 6, pp. 585–593.

[15] J. KEVORKIAN (1966), *Space Mathematics, Part III*, Lectures in Applied Mathematics, vol. 7, Amer. Math. Soc., Providence, R.I.

[16] P. A. LAGERSTROM (1957), *Note on preceding two papers*, J. Math. Mech., 6, pp. 605–606.

[17] ——— (1961), *Méthodes asymptotiques pour l'étude des équations de Navier–Stokes*, Lecture notes, Institut Henri Poincaré, Paris.

[18] ——— (1964), *Laminar Flow Theory*, High Speed Aerodynamics and Jet Propulsion, F. K. Moore, ed., vol. IV, Princeton University Press, Princeton.

[19] P. A. LAGERSTROM AND J. D. COLE (1955), *Examples illustrating expansion procedures for the Navier–Stokes equations*, J. Rational Mech. Anal., 4, pp. 817–882.

[20] G. E. LATTA (1951), *Singular perturbation problems*, Doctoral thesis, California Institute of Technology, Pasadena.

[21] HARTMUT LEGNER (1971), *On optimal coordinates in boundary-layer theory*, Doctoral thesis, Stanford University, Stanford, Calif.

[22] L. PRANDTL (1905), *Über Flüssigkeiten bei sehr kleiner Reibung*, Verh. III Internat. Math. Kongr. (Heidelberg, 1905), Teubner, Leipzig, pp. 484–491; English transl., Tech. Memo. 452, NACA, 1928.

[23] I. PROUDMAN AND J. R. A. PEARSON (1957), *Expansions at small Reynolds numbers for the flow past a sphere and a circular cylinder*, J. Fluid Mech., 2, pp. 237–262.

[24] M. VAN DYKE (1958), *The paraboloid of revolution in subsonic flow*, J. Math. and Phys., 37, pp. 38–51.

[25] ——— (1962), *Higher approximations in boundary-layer theory*, J. Fluid Mech., 14, pp. 161–177 and 481–495.

[26] ——— (1964), *Perturbation Methods in Fluid Mechanics*, Academic Press, New York.

[27] ——— (1964a), *Higher approximations in boundary-layer theory. Part 3: Parabola in uniform stream*, J. Fluid Mech., 19, pp. 145–159.

[28] WOLFGANG WASOW (1965), *Asymptotic Expansions for Ordinary Differential Equations*, Interscience, New York.