ratic equation and what is effectively a linear equation. That is, the root of order ϵ^{-2} was obtained by neglecting the two last terms and writing $3\epsilon^2 x^3 + x^2 \approx 0$, and though this is cubic it has two trivial unacceptable roots $x \approx 0$ (corresponding to the solutions of the quadratic for finite roots) and is therefore equivalent to the linear equation obtained by dividing through by x^2.

The third (also derivative) way systems simplify, often spectacularly, is through the splitting off of autonomous subsystems. By an autonomous subsystem of a system is meant a part of the system (part of the condition together with part of the unknowns) which is complete in itself, i.e. forms an applied mathematical system in its own right, so that it can (in principle, at least) be solved before the rest of the system is considered. The qualifier "autonomous" is by no means superfluous. Thus the system $f(x,y) = 0$, $g(x) = 0$ for the two variables x, y has the autonomous subsystem $g(x) = 0$. It has also the nonautonomous subsystem $f(x,y) = 0$ for y, nonautonomous because it is not definite (well-posed) until x has been determined, which requires the other part of the system.

Systems with autonomous subsystems occur much more often than one may at first realize, since there is an instinctive tendency to concentrate attention on the subsystem and forget that it is part of a larger problem. A particularly contemporaneous illustration of this is provided by the gravitationally determined motion of the sun, a planet, and an artificial satellite; the subsystem of the sun and

planet alone is autonomous, since their motions are unaffected by the satellite and are naturally considered to be given and definite when its motion is under consideration. But there is a very common special kind of system having autonomous subsystems which do not get overlooked just because there are too many of them for any one to be singled out naturally. Such are the initial value problems, which, if well posed for $t_o < t \leq t_1$ with initial conditions at t_o, are also well posed for $t_o < t \leq t_2$ for any t_2 between t_o and t_1, so that the autonomous subsystems constitute a continuous one-parameter family.

For an illustration of the third way of simplifying, note that in HCE theory the five moments ρ, $\underset{\sim}{u}$, p satisfy (in the limit, of course) the autonomous subsystem (4), which is vastly simpler than (1) in having only four independent variables instead of seven. Similarly the "general" (for finite ϵ) pair of simultaneous equations $f(x, y) = 0$, $g(x) + \epsilon\, h(x, y) = 0$ reduces for $\epsilon \to 0$ to the system with an autonomous subsystem considered earlier. The sun-planet subsystem split off only by virtue of the implied limit of (relatively) small satellite mass, as is apparent from the less extreme case of the earth and its natural (rather than artificial) satellite.

The second and third ways both involve a reduction in the number of solutions from which the desired one must be singled out. This is a characteristically asymptotic simplification and, as Friedrichs[1] has affirmed, it justifies the limiting process even though complications arise in other respects. For instance, a linear second order

differential equation may reduce to one of first order but nonlinear. The "number" of solutions must be counted in whatever way is appropriate to the instance: as an integer (e.g. for the polynomical equation); as the dimensionality or number of parameters of a family of solutions (as for an ordinary differential equation); as the dimensionality of a parameter space, or number of independent variables for a function characterizing a solution (as with HCE, where seven reduces to four); or what have you.

In carrying out asymptotic approximations to higher order terms we are aided by the (second) Principle of Recursion, which advises us to treat the nondominant terms as if they were known (even though they involve the unknown solution). The simplified system then determines the unknown in terms of itself, but in an insensitive way suitable (in principle at least) for iterative generation of an asymptotic representation of the solution. This has already been illustrated for one of the finite roots of our cubic equation example. For the numerically large root of (2) we may obtain the recursion formula $x = -(x^2 - \epsilon x - 4)/(3\epsilon^2 x^2)$. However, this is far from unique; by grouping the terms differently we obtain $x = -(x^2 - 4)/(3\epsilon^2 x^2 - \epsilon)$, which is equally suitable, since x has still been solved for from the dominant terms. It would be folly to solve for x from a small term such as ϵx; iteration on $x = (3\epsilon^2 x^3 + x^2 - 4)/\epsilon$ merely produces wilder and wilder ϵ behavior. If one solves from the dominant terms inappropriately, namely in a way which does not give the solution explicitly outright when the small terms are neg-

lected, then one has a scheme which may or may not converge, but which, even if it does, converges at a "finite" rate, not improving the asymptotic order of the solution on each iteration. This is illustrated by putting (2) in the convergent but asymptotically inappropriate recursion form $x = -\left[-(x^2 - \epsilon x - 4)/(3\epsilon^2 x)\right]^{1/2}$, which is quite usable, however, for numerical computation.

This trivial example is so trivial that the emphasis on recursion formulas seems forced. It is true that here and in many many other cases one can simply write down an obvious power series in ϵ and determine the terms order by order. This approach fails, however, whenever a more general representation is required, as is by no means rare. For instance I recently encountered a case where the obvious series needed to be supplemented by a single logrithmic term (which was neither the dominant nor even the next-to-dominant term); the recursion relation generates all the right terms without prejudice as to their form. Generation of terms by recursion is often very clumsy for practical purposes, apart from leading to terms of unexpected form. However, it has a great theoretical advantage when properties of (all terms of) the series are to be derived, since the recursion relation is highly adapted naturally to the use of mathematical induction. (See the final reference for an example.)

The limiting cases we keep referring to are conventionally, in asymptotics, formulated so as to be cases where a parameter (often denoted by λ) approaches infinity. Since I intend asymptotology

to embrace also situations where the limit system itself (not merely arbitrarily near ones) is meaningful (perturbation problems), it is preferable now instead to use a small parameter, conventionally denoted by ϵ ($= 1/\lambda$ for conversion). In fact, it may not be known in advance whether the limit case is meaningful, and, whether or not it is meaningful physically, mathematically it may or may not be so depending on the description employed. This brings us to our third asymptotological principle, the Principle of Interpretation: it is a major task of asymptotological analysis to find variables in which the given problem becomes a perturbation problem (has a meaningful limit situation). This may involve nothing more than recognizing that the original variables are such, as is the case for two roots of the cubic; for the third root, however, the formal limit of (2) is meaningless, but if transformation to the new variable $y = \epsilon^2 x$ is effected first, the equation obtained for y may be solved by perturbation analysis.

The characteristic feature of asymptotic analyses proper, as opposed to perturbation analyses, is the appearance (in both senses) of overdeterminism. Thus the cubic equation (2) with three roots apparently reduces in the limit to a quadratic with only two; the well behaved (for $\epsilon \neq 0$) pair of simultaneous linear equations $x + y = 1$, $x + (1 + \epsilon)y = 0$ formally reduces to a mutually contradictory pair for $\epsilon = 0$; in the initial value problem $\epsilon \frac{d}{dt} z + z = 0$ $(t > 0)$, $z(0) = 1$, for the continuous function $z(t)$, we seemingly have $z(t) = 0$ in the limit, contradicting the initial condition; and the same thing happens in many

less trivial cases (such as the theories of shocks, of boundary layers, and of fast oscillations), as described in detail by Friedrichs.[1] In this connection we have the (fourth) Principle of Wild Behavior, which tells us that apparent overdeterminism arises because (at least some of) the solutions behave wildly in the limit--wildly, that is, compared to our preconceptions, as embodied in the mathematical form of the expressions employed for representing the solutions. Thus in neglecting the cubic (in addition to the linear) term of (3) we have obviously made the implicit assumption that x is not too large (say bounded), which is correct for only two of the roots, while the third behaves "wildly" in becoming infinite (like ϵ^{-2}); the solution of the simultaneous equations is similarly wild (like ϵ^{-1}); the solution of the initial value problem, $z = \exp(-t/\epsilon)$, is wild in having a derivative which, though converging to zero for every fixed positive t, does so nonuniformly and actually becomes infinite for t approaching zero sufficiently rapidly; and similar wildnesses occur in the deeper examples mentioned.

When overdeterminism occurs, if the solution we want is among those still permitted by the formal limit system, well and good: the loss of other solutions is our gain in simplicity (in the second way). If the solution we want is among those lost, then according to the Principle of Wild Behavior we should allow for more general asymptotic behavior of the solution. It is one of the most troublesome difficulties of asymptotological practice to find an appropriate asymptotic form. It is impossible to prescribe a priori all asymptotic representations

that may ever prove useful, but among more general representations to try are two worth specific mention as frequently successful. The first is to supplement the originally expected series with new terms, such as smaller (more negative) powers, as in the case of the cubic equation, or logarithmic ones. The second, effective in many of the deeper problems, including those just referred to (see also a detailed example from my own experience[11]), and illustrated by the initial value problem just exhibited (which may in fact be viewed as an elementary boundary layer problem), is to write the unknown as the exponential of a new unknown represented by a series, the dominant term of which must become infinite (at least somewhere) in the limit if anything is to be gained by so doing.

If there can be overdeterminism there can also be underdeterminism, which means that the original well posed problem reduces formally in the limit to a problem with more than one solution. For instance let A be a known j-by-j matrix, let b and x be j-by-1 matrices, respectively known and unknown, and consider the matrix equation $Ax = b$. Suppose that A and b depend on ϵ and that the determinant of A is zero if and only if $\epsilon = 0$. Then the formal lowest order system $A^{(0)}x^{(0)} = b^{(0)}$ is certainly not well posed. Since $A^{(0)}$ is a singular matrix there exists a 1-by-j matrix n $(\neq 0)$ such that $nA^{(0)} = 0$; for simplicity assume that n is unique (up to a constant factor). If $nb^{(0)} \neq 0$ the limit system obviously has no solution (overdeterminism, as in the previous example of simultaneous linear

equations), so assume $n b^{(0)} = 0$. Then $x^{(0)}$ is not completely determined by the limit system, and we have an example of underdeterminism.

Another excellent and rather typical example of underdeterminism is again the HCE problem. Letting $\lambda \to \infty$ in (1) (after dividing through by λ) leads to the information that $f^{(0)}$ is invariant under collisions, i.e. locally Maxwellian in some (local Galilean) coordinate system, which is very far from determining $f^{(0)}$, since there are five parameters (ρ, $\underset{\sim}{u}$, p) needed to specify such a distribution and we are left unprovided with information on how the parameters at different points of space-time are related. (The Chew-Goldberger-Low[7] theory is another such example.[8]).

In such straits we are rescued by the (fifth) Principle of Annihilation, which instructs us to find a complete set of annihilators of the terms which persist in the limit, apply them to the original system, and then go to the limit after multiplying by an appropriate function of ϵ so that the now dominant terms persist in the limit. By an annihilator of a mathematical entity is meant an operator which results in zero when applied to the entity. (Of course there are complicated cases in which this produces only some of the missing information, and the same procedure must be reapplied, perhaps repeatedly.).

In the matrix example, the terms $A^{(0)}x^{(0)}$ and $b^{(0)}$ which persist in the limit are annihilated by multiplication on the left by n. Applying this annihilator to the original equation, dividing by ϵ, and taking the limit gives what may be written

$$\lim_{\epsilon = 0} \left\{ \epsilon^{-1} n [A - A^{(0)}] \right\} x^{(0)} = \lim_{\epsilon = 0} \left\{ \epsilon^{-1} n [b - b^{(0)}] \right\} \tag{5}$$

or $nA^{(1)}x^{(0)} = nb^{(1)}$ if A and b are expandable in integral powers of ϵ. In the normal case this provides just the one extra condition needed to determine $x^{(0)}$, which by the condition $A^{(0)}x^{(0)} = b^{(0)}$ was determined only up to a solution p of $A^{(0)}p = 0$. In the abnormal case that (5) is not an independent condition, there is a linear combination of $A^{(0)}x^{(0)} = b^{(0)}$ and (5) which gives $0 = 0$. The formation of this linear combination is then our new annihilator, the application of which to $Ax = b$ and $\epsilon^{-1} n [A - A^{(0)}] x = \epsilon^{-1} n [b - b^{(0)}]$ leads to a new extra condition which will normally be independent and provide the missing piece of information.

In the HCE problem there are five scalars (mass, three components of momentum, and energy) which are preserved by collisions, so that taking the corresponding moments of (1) annihilates the right side. These are therefore annihilators of the dominant terms, which is why they are applied to (1) to obtain the five hydrodynamic equations relating the values of ρ, $\underset{\sim}{u}$, p (and therefore f which is expressed in terms of them) at different points of space-time.

It is through the application of the Principle of Annihilation that the Principle of Simplification is maintained. The loss of solutions in a limit simplifies a system, while the gain of solutions, or loss of