PRENTICE-HALL INTERNATIONAL, INC. London
PRENTICE-HALL OF AUSTRALIA, PTY., LTD. Sydney
PRENTICE-HALL OF CANADA, LTD. Toronto
PRENTICE-HALL FRANCE, S.A.R.L. Paris
PRENTICE-HALL OF JAPAN, INC. Tokyo
PRENTICE-HALL DE MEXICO, S.A. Mexico City

Mathematical Models in Physical Sciences

Proceedings of the Conference at the University of Notre Dame, 1962

Edited by
Stefan Drobot
Professor of Mathematics
University of Notre Dame
and
PAUL A. VIEBROCK
University of Notre Dame

PRENTICE-HALL, INC. Englewood Cliffs, N. J.

Preface

On April 15-17, 1962, a conference on "Mathematical Models in Physical Sciences" was held at the University of Notre Dame. The conference was sponsored by the National Science Foundation under grant No. NSF-G23332.

I wish to express my appreciation to my graduate assistant, Paul A. Viebrock, whose collaboration in compiling and editing made the publication of these proceedings possible.

Stefan Drobot

Editor

University of Notre Dame

February, 1963

They would, presumably, have some physical validity under conditions when they admit solutions at all.

Perhaps the greatest significance of the guiding-center plasma model is that it turns out to be partly macroscopic. Thus there are many important problems in which adherence to a purely microscopic formulation offers no advantage over a macroscopic picture; this despite the much greater computational complexity which results from the detailed pursuit of individual particles.

Both the mathematical and physical significance of the guiding-center fluid and plasma systems remains largely to be discovered.

REFERENCES*

1. N. Bogolubov, J. Phys. USSR **10**, 265 (1946).
2. K. O. Friedrichs, Willard Gibbs lecture, Bull. Am. Math. Soc. **61**, No. 6, 485-504 (1955).
3. Harold Grad
   1958, "Principles of the Kinetic Theory of Gases", in Handbuch der Physik, Vol. XII, Springer-Verlag.
   1960, J. Chem. Phys. **33**, p. 1342.
   1961a, Comm. Pure and Appl. Math. **14**, p. 323.
   1961b, "Microscopic and Macroscopic Models in Plasma Physics" in Proceedings of the Symposium on Electromagnetics and Fluid Dynamics of Gaseous Plasma, Polytechnic Institute of Brooklyn.
   1963a, Phys. Fluids **6**, p. 147.
   1963b, "Asymptotic Theory of the Boltzmann Equation—II", to appear in the Proceedings of the Third International Conference on Rarefied Gases, Paris, 1962.
4. D. Hilbert, Math. Ann. **72**, 562 (1912); also "Grandzüge einer Allgemeinen Theorie der Linearen Integralgleichungen", Chelsea Publ. Co., New York, 1953.
5. R. F. Probstein and Y. S. Pan, "Shock Waves and the Leading Edge Problem", to appear in the Proceedings of the Third International Conference on Rarefied Gases, Paris, 1962.

*More complete references for the material in Sections 1, 2, and 3 are cited in the references, Grad 1963a, 1958, and 1961b, respectively.

*Asymptotology**

MARTIN D. KRUSKAL
Plasma Physics Laboratory
Princeton University

When I first saw the program for this conference I was mildly curious about why my talk was scheduled at the end of the first session, following the opening lecture by Professor Grad. Although accepted conference manners (conventional convention conventions, I almost said) forbade inquiring of our genial organizers, I now know the reason—Harold's stimulating and excellent lecture has roused a furor of excitement and even controversy, as they must have foreseen, and it is my function to calm you down, bore you perhaps, and send you off properly soothed and relaxed to enjoy tonight's banquet.

The subject of this conference is unusual, and if I am not at all confident that my chosen topic is entirely appropriate, I am em-

*This work was supported under Contract AT(30-1)-1238 with the Atomic Energy Commission.

boldened to proceed because of a conviction that it would be out of place anywhere else. But I do feel some trepidation at having Professor Friedrichs in the audience, since I am so heavily indebted to his most enlightening 1955 Gibbs Lecture article,[1] already referred to by Grad.

Asymptotics is the science which deals with such questions as the asymptotic evaluation of integrals, of solutions of differential equations, etc., in various limiting cases. Elements of this science may be learned from the works of van der Corput,[2] Erdélyi,[3] and de Bruijn,[4] and advanced aspects from the numerous references in Friedrichs' cited article. By asymptotology I mean something much broader than asymptotics, but including it; pending further elaboration, I would briefly define asymptotology as the art of dealing with applied mathematical systems in limiting cases.

The first point to note here is that asymptotology is an art, at best a quasi-science, but not a science. Indeed, this explains much of my difficulty both in expounding my material and in finding an appropriate occasion to do so, and it may serve handily to excuse my effort for lacking the high degree of polish which Dean Rossini in his opening remarks assured us we may expect of the presentations (and indeed there does seem to be much Polish about this conference). It explains, too, why I am unable to support the corpus of my dissertation with the hard bones of theorems but must be content with a cartilage of principles, into seven of which I have distilled

whatever of asymptotology I have been able to formulate appropriately and sufficiently succinctly.

The aspect of the definition of asymptotology just given which is most in need of explanation is the concept of applied mathematical system. An applied mathematical system is merely the mathematical description of a physical (or occasionally biological or other) system in which the variables expressing the state of the system are complete. The importance of formulating problems in terms of complete state variables constitutes a preliminary principle, not particularly of asymptotology but of applied mathematics in general, the Principle of Classification (or, perhaps better, of Determinism). It is illustrated by the overpowering tendency, in treating classical mechanical problems, to enlarge the configuration space to a phase space, since the phase (configuration together with its rate of change) but not the configuration alone constitutes a complete description of a classical mechanical system. Consider also the tendency, in treating probabilistic mechanical problems, to switch over from this original description, which is incomplete because, for instance, the mechanical "state" at one time does not determine the "state" at another time, to a new description in terms of a probability distribution function of the old "states," which function evolves "deterministically" in time and is therefore preferable as a state description. This Principle is obviously closely related to the notion of a well posed problem emphasized by

Hadamard. Its particular relevance to asymptotology comes about because only after one has singled out ("determined") an individual solution (or completely "classified" the family of solutions) can one reasonably inquire into its asymptotic behavior.

Asymptotology is important because the examination of limiting cases seems to be the only satisfactory effective method of proceeding with the analysis of complicated problems (systems) when exact mathematical methods are of no (further) avail (and is often preferable even when they are). It is of value both for obtaining qualitative information (insight) about the behavior of a system and its solutions and for obtaining detailed quantitative (numerical) results. Thus it is hardly surprising that examples, from trivial ones to the most profound, are found everywhere throughout the fields to which analysis (in the technical sense as a branch of mathematics) is applied.

An excellent example of asymptotology is the familiar Hilbert[5] or Chapman-Enskog[6] ("HCE" from now on) theory of a gas described by the Boltzmann equation

$$(1) \qquad \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \mathbf{a} \cdot \frac{\partial f}{\partial \mathbf{v}} = \lambda \int d^3v \, d\Omega \, |\mathbf{v} - \mathbf{v}'| \, \sigma [\bar{f}\bar{f}' - ff']$$

in the limit of high density ($f \to \infty$) or equivalently of frequent collisions ($\lambda \to \infty$), which Grad has already discussed this afternoon. Another example is the Chew-Goldberger-Low[7] theory of the so-called Vlasov[8] system of equations governing an ideal collisionless plasma

and its electromagnetic field in what is often called the strong magnetic field (or small gyration radius) limit but is formally best treated[9] as the limit of large particle charges. In the general theory of relativity there is the fundamental Einstein-Infeld-Hoffman[10] derivation of the equation of motion of a "test particle" (one not influencing the space-time metric, i.e., one of negligible mass) by treating it (its world-line, rather) as an appropriate singularity in the metric and letting the strength of the singularity approach zero. Hydrodynamics is rich in asymptotology (theory of shocks as arising in the limit of small viscosity and heat conductivity, theories of strong shocks and of weak shocks, shallow water theory, and so on and on), and so is elasticity. Kirchoff's laws for electrical circuits can be properly derived from Maxwell's equations only by going to the limit of infinitely thin conductors (wires). Simpler examples also abound and are encountered daily by the practicing applied mathematician and theoretical physicist. Naturally it is not practical to discuss deep examples in detail here, so I shall have to confine myself to brief remarks about them, relying for illustration mainly on simple and often trivial instances.

It should now be apparent, I hope, that whatever features such important, wide-spread, and diverse examples may have in common, and whatever lessons for future application may be gleaned from studying them, are well worth formulating and eventually standardizing. Even the many (most? far from all, as I know from my ac-

quaintance) applied mathematicians (etc.) who have become familiar by experience with asymptotological principles, at least in the sense of knowing how to apply them in practice,—even they must inevitably benefit from the introduction of a standard terminology and of the clarity of expression it permits. Implicit knowledge, no matter how widely distributed, deserves explicit formulation, but I am aware of no efforts in this direction which attempt to go anything like so far as I am doing here, though there are some related suggestions in Friedrichs' article.

The final possible obscurity in our previous tentative definition of asymptotology is what it means to deal with a system. To clarify this, we might alternatively define asymptotology as the art of describing the behavior of a specified solution (or family of solutions) of a system in a limiting case. And the answer quite generally has the form of a new system (well posed problem) for the solution to satisfy, although this is sometimes obscured because the new system is so easily solved that one is led directly to the solution without noticing the intermediate step.

To illustrate first by a trivial example, suppose it is desired to follow the (algebraically) largest root x of the simple polynomial equation

$$3\varepsilon^2x^3 + x^2 - \varepsilon x - 4 = 0 \tag{2}$$

in the limit $\varepsilon \to 0$. There is one root of order ε^{-2} obtained by treating the first two terms as dominant, $x \approx -\frac{1}{3}\varepsilon^{-2}$, for which indeed

the other two terms are relatively negligible (even though one of them is absolutely large, of order ε^{-1}), but which is negative. The other two roots are finite, obtained by neglecting the terms with ε factors, $x \approx \pm 2$, the one sought having the plus sign. If we desire it to higher order, incidentally, we may put (2) for this root in the "recursion" form

$$x = 2\left(1 - \frac{3}{4}\varepsilon^2x^3 + \frac{1}{4}\varepsilon x\right)^{1/2}, \tag{3}$$

expand out the right side in powers of ε, and generate better and better approximations for x by continually substituting the previously best approximation into the right side. But this is irrelevant to the present point, which is that (the problem of the algebraically largest root of) the original cubic equation (2) has been replaced by (the problem of the algebraically largest root of) the quadratic equation $x^2 - 4 \approx 0$, or more exactly $x^2 - (4 - 3\varepsilon^2x^3 + \varepsilon x) = 0$, the quantity in parentheses being treated as known.

In the HCE treatment of system (1) in the limit $\lambda \to \infty$, the original integro-differential equation in the seven independent variables t, $\mathbf{x}$, $\mathbf{v}$ gets replaced by the set of coupled partial differential (hydrodynamic) equations

$$\begin{aligned} &\frac{\partial \rho}{\partial t} \approx -\frac{\partial}{\partial \mathbf{x}} \cdot (\rho \mathbf{u}), \\ &\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \frac{\partial}{\partial \mathbf{x}}\mathbf{u} \approx -\frac{1}{\rho}\frac{\partial p}{\partial x}, \\ &\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \frac{\partial}{\partial \mathbf{x}}\right)(\rho^{-5/3}p) \approx 0 \end{aligned} \tag{4}$$

in the four independent variables t, x; here ρ, u, p are of course the usual velocity space moments of f.

These examples clearly illustrate the first asymptotological principle, which is in fact largely the raison d'être of asymptotology. This Principle of Simplification states that an asymptotological (limiting) analysis tends to simplify the system considered. This can occur in at least three general ways.

The basic way systems simplify is merely by the neglect of terms (or, in higher order analyses, at least treatment of small terms as if known, as in the case of the cubic equation earlier). Thus the polynomial equations $x^5 - \varepsilon x + 1 = 0$ and $x^6 + ax^4 + \varepsilon x^3 + 1 = 0$, without getting lower in degree as the cubic did, nevertheless become simple enough in the limit $\varepsilon \to 0$ to be explicitly solvable algebraically. Differential equations in irregular domains approximating regular ones may in the limit become solvable by separation of variables. In other cases the coefficients may become so simple in the limit as to permit solution by Fourier or other transform. These are typical instances of perturbation theory; there are of course also many instances where the simplification which occurs does not appreciably facilitate the further analysis of the system.

A derivative way in which systems simplify, sometimes striking in effect, is the decomposition of the system into two or more independent systems among which the solutions are divided, so that the

particular solution of interest satisfies a system with fewer solutions and hence usually in some sense of lower order. Thus the cubic polynomial equation considered earlier split up into a quadratic 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 $3\varepsilon^2x^3 + x^2 \approx 0$, 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 conditions 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 $t_0 < t \leq t_1$ with initial conditions at t_0, are also well posed for $t_0 < t \leq t_2$ for any t_2 between t_0 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 $\rho, \mathbf{u}, p$ 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 $f(x, y) = 0$, $g(x) + \varepsilon h(x, y) = 0$ reduces for $\varepsilon \to 0$ 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 polynomial 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 of 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 $x = -(x^2 - \varepsilon x - 4)/(3\varepsilon^2 x^2)$. However, this is far from unique; by grouping the terms differently we obtain $x = -(x^2 - 4)/(3\varepsilon^2 x^2 - \varepsilon)$, 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 $x = (3\varepsilon^2 x^3 + x^2 - 4)/\varepsilon$ 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 neglected, 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 in each iteration. This is illustrated by putting (2) in the convergent but asymptotically inappropriate recursion form $x = -[-(x^2 - \varepsilon x - 4)/(3\varepsilon^3 x)]^{1/2}$, 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 logarithmic 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 de-

rived, 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 $y = \varepsilon^2 x$ 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 $\varepsilon \neq 0$) pair of simultaneous linear equations $x + y = 1$, $x + (1 + \varepsilon)y = 0$ formally reduces to a mutually contradictory pair for $\varepsilon = 0$; in the initial value problem $\varepsilon \frac{d}{dt} z + z = 0$ $(t > 0)$, $z(0) = 1$, 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, $z = \exp(-t/\varepsilon)$, 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 $\varepsilon = 0$. 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 $nb^{(0)} = 0$. 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, **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

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

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 $\varepsilon^{-1}n[A - A^{(0)}]x = \varepsilon^{-1}n[b - b^{(0)}]$ 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 compo-

nents 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, $\mathbf{u}$, 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 information,* would "complicate" it if we were not able to recover sufficient additional conditions to make up for the information lost.

The basic way systems simplify is by the neglect of terms, as stated earlier. But it commonly happens that the relative asymptotic magnitude of two terms to be compared depends upon some knowledge not yet available or on some assumption or decision not yet made. According to the (sixth) Principle of Maximal Balance (or of Maximal Complication†), for maximal flexibility and generality we should keep both terms, i.e., we should allow for the possibility or assume that they are comparable. In the case of incomplete knowledge this is mere prudence; any term in an equation

*Use of this terminology is justified even from the technical viewpoint of information theory, suggesting the possibility of assigning a measure to the decrease in the number of solutions occurring in a limit.

†Partly as a consequence of Professor Friedrichs' comment at the conclusion of my lecture, I now feel that "Minimal Simplification" is more appropriate here.

definitely smaller in order of magnitude than another term may be considered negligible, but no term should be neglected without a good reason. In the case of a pending assumption or decision, the desire to balance two such competing terms helps to determine the choice.

The most widely applicable and hence most informative ordering is that which simplifies the least, maintaining a maximal set of comparable terms. Quite often there is more than one possible maximal set of terms, with no set including all terms of any other. (Sets of terms form a lattice ordered by inclusion.) Each maximal set corresponds to different asymptotic behavior. The solutions may split up according to which behavior they have (second way of simplifying), as with the cubic, or each solution may exhibit a variety of different behaviors, in different regions, as with a boundary layer phenomenon.

For instance in the case of the cubic equation, how could we know that two solutions are finite and one of order ε^{-2}? Put another way, why did we not assume the first and third terms to be the dominant ones, or the second and third, or so on? In this particular case there is an easy answer: if we had, we would have obtained a "solution" for which the neglected terms were not in fact negligible compared to the supposed dominant terms, i.e., the "solution" found would not have been self-consistent. But suppose there were several more terms, would we have had to try every pair? (Or suppose there were two independent small parameters δ and ε in-

stead of only one.) Clearly, no matter which terms are dominant x will behave predominantly as some power of ε. We therefore assume the general representation $x \approx a\varepsilon^q$ and wonder what value of q to take. One might in fact choose arbitrarily any value for q but will then generally find that for finite a only one term of (2) dominates, which is nonsensical, so that $a = \infty$ (if it was the constant term), which is not legitimate, or else $a = 0$ (if it was one of the others), which, if more legitimate, is certainly no more useful. A value of q will only be "proper" if we end up with a representation which is "maximally complicated" in that it really consists of one term $a\varepsilon^q$ instead of "no terms" such as 0 or ∞. If we put $x \approx a\varepsilon^q$ into (2) the successive terms vary as ε to the respective powers $3q + 2$, $2q$, $q + 1$, 0, and it is easy to see that only $q = 0$ or $q = -2$ make two (or more) powers equal minima.

On the side it might be of interest to mention a graphical method of finding the proper values of q which apparently goes back to Newton. It is hardly needed in the present simple illustration but can be a great time-saver in more involved examples (also those of higher dimensionality). We plot each term of (2) as a point on a graph, the abscissa being the exponent of x and the ordinate that of ε (see four heavy points in Figure 1); the coefficient is ignored so long as it is not zero. The specification of a definite relationship between x and ε (i.e. of a definite value of q) leads to the identification of the asymptotic behavior of all terms (present or not) cor-

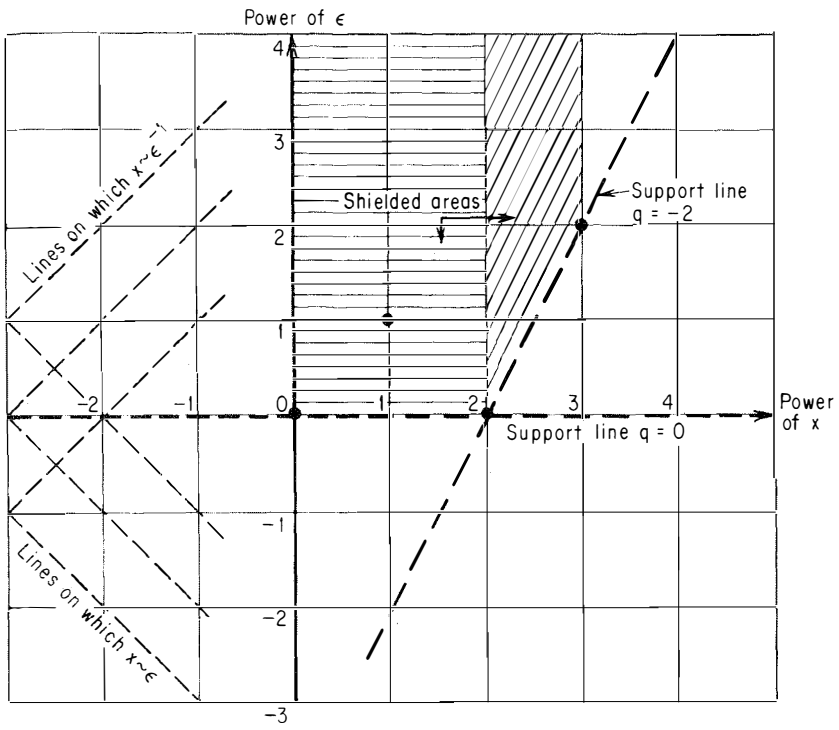

Figure 1

responding to points which are on a common line with a definite slope. Thus, for $x \sim \varepsilon$ all points on the same down-slanting (from left to right) 45° line correspond to a common asymptotic behavior, while for $x \sim \varepsilon^{-1}$ the same holds for up-slanting 45° lines (see dotted lines). Since the smaller the power of ε the larger the term, we seek lines passing through (at least) two graphed points and having no graphed points below them. We may think of finding the lower convex support lines of the set of graphed points, perhaps

kinesthetically by imagining pushing a line up from below until it first hits a graphed point and then rotating it around that point until it next hits a second graphed point. It is immediately apparent from Figure 1 that there are just two such lines and that they correspond to $q = 0$ and $q = -2$ (see heavy dashed lines). It is also clear that the point (1, 1), like all points in a semi-infinite vertical strip (see horizontally shaded area), are "shielded" by the points (0, 0) and (2, 0) and can never be on a support line; it is indeed obvious that εx is negligible with respect to either x^2 or 4 no matter how x varies with ε. Similarly there is a semi-infinite vertical strip shielded by the points (2, 0) and (3, 2) (see diagonally shaded area). In more complicated cases we can thus exclude terms wholesale from competition.

To return to our proper business, illustration of the Principle of Maximal Complication, consider the problem of finding the lowest frequency of vibration and the corresponding form of vibration of a uniform membrane stretched between two close wires lying in a plane, one of which we take straight for simplicity. The equation for the standing vibration of a membrane is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \nu^2 u = 0, \tag{6}$$

where u is the displacement normal to the (x, y) plane, which is the rest plane of the membrane (the plane containing the wires), and ν is the frequency of vibration of the mode. Let the equations of the wires in the (x, y) plane be $y = 0$ and $y = \varepsilon Y(x)$, where ε of course

is the small parameter of closeness. We may suppose $Y(x_1) = Y(x_2) = 0$ so as to have to consider only the finite region $x_1 < x < x_2$, $0 < y < \varepsilon Y(x)$. Imposing the condition $u = 0$ on the boundary of this region and (6) inside the region, we have an eigenvalue problem for the lowest eigenvalue ν and its corresponding eigenfunction u. This is one common type of asymptotic problem, asymptotic rather than "perturbational" in that there is no limit problem because the region of interest disappears in the limit. The remedy for this is well known;[1] we rescale the variables appropriately, in this case introducing $\eta \equiv \varepsilon^{-1} y$ so that the region in the (x, η) plane becomes $x_1 < x < x_2$, $0 < \eta < Y(x)$, and (6) becomes

$$\frac{\partial^2 u}{\partial x^2} + \varepsilon^{-2} \frac{\partial^2 u}{\partial \eta^2} + \nu^2 u = 0. \tag{7}$$

Taking the asymptotic behavior of each term at its face value (but remembering that ν is not yet determined), we deem the first term negligible compared to the second, and (by the Principle) assume $\nu^2 \sim \varepsilon^{-2}$ to balance the second and third terms. Introducing $\omega \equiv \varepsilon \nu$ we write (7) as

$$\frac{\partial^2 u}{\partial \eta^2} + \omega^2 u = -\varepsilon^2 \frac{\partial^2 u}{\partial x^2}. \tag{8}$$

To lowest order we neglect the right side of (8), whereupon x degenerates from an independent variable to a mere parameter. The really proper treatment at this point, by the Principle of Recursion, would be to treat the right side of (8) as known, solve for u on the left in the form of an integral representation (involving the simple,

well known, explicit Green's function), and try to obtain u iteratively. Instead we shall do something similar but simpler, more or less paralleling the lowest order version of the proper treatment. For each x we have, to lowest order, a simple eigenvalue problem with the lowest eigenstate $u = A \sin(\pi\eta/Y)$ and eigenvalue $\omega = \pi/Y$. But ω so defined depends on x, which is impermissible, so we take A(x) to be a Dirac delta function, the location of whose singularity we take to be at the maximum of Y(x) in order to have the smallest ω; for simplicity we assume the maximum of Y to be unique and to occur at x = 0. In a sense we have now solved the problem originally posed, but since our answer is singular it is not entirely satisfactory (see the next and final Principle to be formulated). Indeed, since our "solution" is singular in its x dependence, we ought to worry whether our earlier neglect of $\varepsilon^2 \frac{\partial^2 u}{\partial x^2}$ was justified, and we might well be curious anyway about the true detailed x dependence which we have cavalierly expressed as a delta function. Since the significant behavior occurs near x = 0 we introduce $\xi = \delta^{-1}x$, where δ is a small parameter to be determined (related to ε). We also write $\omega = \omega_0 + \hat{\omega}$, where $\omega_0 = \pi/Y(0)$ and $\hat{\omega}$ is small. Since $\frac{\partial^2 u}{\partial \eta^2} \approx -\pi^2 Y(x)^{-2} u$, from (8) we obtain

$$(9) \qquad \left[\frac{\pi^2}{Y(\delta\xi)^2} - \omega^2\right] A \approx \frac{\varepsilon^2}{\delta^2} \frac{d^2A}{d\xi^2}.$$

Let $Y(\delta\xi) = Y(0) + \frac{1}{2} Y''(0) \delta^2 \xi^2 + \ldots$ with $Y''(0) < 0$, whereupon this becomes

$$(10) \qquad \left[-\frac{\pi^2 Y''(0)}{Y(0)^3} \delta^2 \xi^2 - 2\omega_0 \hat{\omega}\right] A \approx \frac{\varepsilon^2}{\delta^2} \frac{d^2A}{d\xi^2}.$$

According to the Principle of Maximal Complication we choose the as yet undetermined asymptotic behaviors so as to keep all the terms in the equation and are thus led to take $\delta = \varepsilon^{1/2}$ and $\tilde{\omega} = \varepsilon^{-1}\hat{\omega}$, obtaining

$$(11) \qquad \frac{d^2A}{d\xi^2} + \frac{\pi}{Y(0)} \left[\frac{\pi Y''(0)}{Y(0)^2} \xi^2 + 2\tilde{\omega}\right] A \approx 0.$$

On the ξ distance scale A must vanish at "infinity," and we have a well known eigenvalue problem arising in the quantum theory of the harmonic oscillator. The lowest eigenfunction is the Gaussian $A = \exp\left\{-\frac{\pi}{2} Y(0)^{-3/2} [-Y''(0)]^{1/2} \xi^2\right\}$ with real eigenvalue $\tilde{\omega} = \frac{1}{2}[-Y''(0)/Y(0)]^{1/2}$.

Incidentally, if we should be interested in the behavior of u for $|x|$ not very small, where u decreases rapidly, a different procedure must be used. The right side of (8) cannot be neglected there, since $\omega \approx \pi/Y(0)$ does not even approximate the local eigenvalue $\pi/Y(x)$ for which the left side can vanish with $u \neq 0$. The device mentioned earlier of representing the unknown as an exponential works here; with $u = \exp v$, (8) becomes

$$(12) \qquad \frac{\partial^2 v}{\partial \eta^2} + \left(\frac{\partial v}{\partial \eta}\right)^2 + \omega^2 = -\varepsilon^2 \left[\frac{\partial^2 v}{\partial x^2} + \left(\frac{\partial v}{\partial x}\right)^2\right].$$

We may assume that v is expandable as a series in ε, $v = \varepsilon^{-1}[v^{(0)} + \varepsilon v^{(1)} + \ldots]$, where the leading term has been taken large of order ε^{-1} to permit the right side of (12) to contribute. We must have

$\frac{\partial v^{(0)}}{\partial \eta} = 0$ or the left side will dominate again, so $v^{(0)}$ is a function of x only, and to dominant terms (12) becomes

$$\frac{\partial^2 v^{(1)}}{\partial \eta^2} + \left(\frac{\partial v^{(1)}}{\partial \eta}\right)^2 + \omega_0^2 = -\left(\frac{\partial v^{(0)}}{\partial x}\right)^2.$$

Viewed as an equation for $v^{(1)}$ this can be linearized and "homogenized" by reversing the exponentiation procedure, namely by introducing $w = \exp v^{(1)}$, whence

$$\frac{\partial^2 w}{\partial \eta^2} + \left[\omega_0^2 + \left(\frac{\partial v^{(0)}}{\partial x}\right)^2\right] w = 0.$$

Together with the boundary conditions on w (that it vanish at $\eta = 0$, Y(x)) this is an eigenvalue problem which determines the variation of $v^{(0)}$,

$$\omega_0^2 + \left(\frac{\partial v^{(0)}}{\partial x}\right)^2 = [\pi / Y(x)]^2,$$

as well as the η dependence of w (sinusoidal). All that the device has amounted to in this case, of course, is factoring out (from u) a fast varying function of x, but the use of the exponential representation has led to that procedure in a natural and systematic way.

We complete our list with the simple (seventh) Principle of Mathematical Nonsense: if, in the course of an asymptotological analysis, a mathematically nonsensical expression appears, this indicates that the asymptotology has not been done correctly or at least not carried out fully (although even incomplete it may be satisfactory for one's purposes). One may come upon expressions such as 0/0, divergent sums or integrals, singular functions, etc.,

and whether they are to be considered nonsensical sometimes depends on the use they are to be put to. In the just discussed membrane vibration problem the first instance of mathematical nonsense was the disappearance in the limit of the region over which the partial differential equation was to be solved, the second was perhaps the dependence of ω on x, and the third was the response to this, the use of a singular (delta) function.

Frequent in asymptotological analyses is the occurrence of phenomena on different scales of distance or time. The HCE problem is a well-known case (as Grad has just pointed out), since if f is not prescribed Maxwellian at the initial instant, there is a relatively short period of time (the order of a collision time) during which f becomes Maxwellian, while the five moments remain approximately constant, and a relatively long period (of order λ times as long) during which the five moments (hydrodynamic variables) vary but f maintains its Maxwellian form. For an extremely simple example of the same type, consider the familiar electric circuit equation $V = RI + L\dot{I}$, where the voltage V(t) is an imposed function of time, the current I(t) is to be found, the resistance R and the inductance L are positive constants, and we choose to examine the limit $L \to 0$. Treating $L\dot{I}$ as if it were known, we immediately obtain a recursion formula for I,

$$\begin{aligned} I &= \frac{1}{R}(V - L\dot{I}) \\ &= \frac{1}{R}\left[V - \frac{L}{R}\dot{V} + \left(\frac{L}{R}\right)^2 \ddot{V} - \left(\frac{L}{R}\right)^3 \dddot{V} + \ldots\right], \end{aligned} \tag{13}$$

which is fine except for not in general satisfying the arbitrary initial condition on I natural for the original first order differential equation. For short times (of order L) $\dot{I}$ is large and V approximately constant, so that the difference of I from its quasi-equilibrium value V/R decays like $\exp(-Rt/L)$; after this transient has died out (13) holds. Incidentally, the expression in brackets in (13) is just like the Taylor expansion in powers of L of V evaluated at the argument $t - L/R$ except for a factor of $(n-1)!$ in the denominator of the n-th term, which shows that the asymptotic series (13) for I cannot be expected to converge even if V is analytic (which does not stop it from being very useful).

In phenomena with behavior on two different time scales there is a widely pertinent distinction to be observed between finite conservative systems on the one hand and infinite or dissipative systems on the other. For instance, the well-known problem of the harmonic oscillator with slowly varying coefficient of restitution,[12] $\ddot{x} + k(\varepsilon t)x = 0$, is an example of the first kind; on the short (finite) time scale k is approximately constant and the oscillator simply oscillates steadily, while on the long ($\sim\varepsilon^{-1}$) time scale the frequency and amplitude of the oscillation vary in response to the variation in k. Contrast with this the behavior of the dissipative electric circuit, where only initially the current I varies on the short time scale, swooping toward its quasi-steady value. The HCE example shows that a conservative system can act the same way so long as

it is infinite; in this case the decay comes about by a process of "phase mixing," and is possible because the Poincaré recurrence time is infinite.

The asymptotic separation of time scales is the basis for an exciting recent approach in statistical mechanics.[13] Typically one obtains equations for the one-particle and the two-particle distribution functions f_1 and f_2 for a gas of appropriate characteristics, and finds that f_1 can vary only slowly, but that f_2 can vary quickly so as to phase-mix towards a quasi-steady distribution as t gets large on the short time scale while remaining small on the long time scale. The limiting distribution f_2 is a functional of f_1, which when substituted into the equation for $\dot{f}_1$ leads to a closed "kinetic equation" for f_1. The irreversibility (timewise) of this kinetic equation comes about in a natural way, in that the limiting f_2 depends on which direction t is taken to the limit (on the short time scale), whether to plus or to minus infinity. It is a major triumph of this approach that the "Stosszahlansatz" can for the first time be actually derived (under moderate smoothness assumptions).

To return to the finite case, I am glad to take the opportunity of advertising a recent paper[14] in which I have elaborately worked out the asymptotic theory of finite systems of ordinary differential equations depending on a small parameter ε which to lowest order have all solutions periodic. Applied to Hamiltonian systems the theory leads to the existence of adiabatic invariants which are constant (integrals) to all orders in ε.

We are all familiar with those rather unsatisfactory research papers in which the author makes a series of largely arbitrary ad hoc approximations throughout, often dubious without (sometimes even with) the author's intuitive grasp of the situation. These "ad-hoaxes" have their place and utility, but how much more desirable and convincing is a properly worked out and elegant asymptotological treatment, with any arbitrary assumptions (like remarkable coincidences in a well constructed mystery story) made openly and aboveboard right at the beginning where anyone can assess their merits for himself, and with the later development unfolding naturally and inexorably once a definite problem and the limit in which it is to be considered have been settled upon!

The art of asymptotology lies partly in choosing fruitful limiting cases to examine—fruitful first in that the system is significantly simplified and second in that the results are qualitatively enlightening or quantitatively descriptive. It is also an art to construct an appropriate generic description for the asymptotic behavior of the solution desired. The scientific element in asymptotology resides in the nonarbitrariness of the asymptotic behavior and of its description, once the limiting case has been decided upon.

Molière has one of his characters observe that for more than forty years he has been talking prose without knowing it. It is doubtful that he benefited from the discovery, but I hope that you will be more fortunate and not disappointed in having by now discovered that asymptotology is what you have been practicing all along!

REFERENCES

1. K. O. Friedrichs, Bull. Amer. Math. Soc. **61**, 485 (1955).
2. van der Corput, references given in Introduction of reference 3.
3. A. Erdélyi, "Asymptotic Expansions," Dover Publ. (1956).
4. N. G. de Bruijn, "Asymptotic Methods in Analysis," North-Holland Publ. (1958).
5. D. Hilbert, Math. Ann. **72**, 562 (1912).
6. S. Chapman and T. G. Cowling, "The Mathematical Theory of Non-uniform Gases," Cambridge [Eng.] Univ. Press (1952).
7. G. Chew, M. Goldberger, and F. Low, Proc. Roy. Soc. A236, 112 (1956).
8. A. Vlasov, J. Phys. (U.S.S.R.) **9**, 25 (1945).
9. M. Kruskal, in "La théorie des gaz neutres et ionisés," edited by C. De Witt and J. F. Detoeuf, Hermann, Paris, and John Wiley, New York (1960).
10. A. Einstein, L. Infeld, and B. Hoffman, Ann. Math. **39**, 65 (1938).
11. M. Kruskal, Rendiconti del Terzo Congresso Internazionale sui Fenomeni d'Ionizzazione nei Gas tenuto a Venezia, Societá Italiana di Fisica, Milan (1957), same as U.S. Atomic Energy Commission Report No. NYO-7903 (1958).
12. R. Kulsrud, Phys. Rev. **106**, 205 (1957).
13. E. Frieman, J. Math. Phys. (March, 1963), in press.
14. M. Kruskal, J. Math. Phys. 3, 806 (1962).

DISCUSSION

Friedrichs: I would like to make a few comments concerning the terminology of these various principles. In your "Principle of Maximal Complication" the term "complication" doesn't quite convince me. You want to catch a wild solution. You have to tamper

with it. If you tamper too much it tends to something trivial. So you have to make it tame but you have to minimize the tameness.

I wanted to use the term "complication" in connection with the "Principle of Simplification" because so many complications involved in the mathematical formulation of problems in physics are thus due to simplification. For example, if you take a second order equation with ε in front of the second derivative, for which the limit equation for $\varepsilon \to 0$ is of the first order, you don't know whether the boundary condition gets lost, or the new equation is nonlinear; it might have a singularity or the solution may not be single-valued: There are lots of complications. But in spite of this you want to make the simplification. So perhaps you should call this the "Principle of Simplification in Spite of the Resulting Complications."

Kruskal: Obviously all these things require interpretation of the words. It is hard to get labels that are entirely satisfactory.

Models of Total Ignorance in Quantum Mechanics

FREEMAN J. DYSON
Institute for Advanced Study
Princeton, New Jersey

1. DEFINITION OF ENSEMBLES

The theory I will talk about is a new kind of statistical mechanics, invented about 10 years ago by Wigner. In ordinary statistical mechanics we assume that we are totally ignorant of the state of a system. We then deduce properties of the system which hold on the average, where the average is defined with respect to a suitably large ensemble of possible states. In the new statistical mechanics we assume that we are ignorant not only of the state of a system, but also of the nature of the system. We then deduce properties which hold on the average, where the average must be defined in terms of an ensemble of systems.

A system is described in quantum mechanics by a finite matrix M operating in a finite vector-space H. We assume that all we