

ORDINARY DIFFERENTIAL EQUATIONS

W. A. COPPEL

(Received 25 May 1976)

This is the text of an invited lecture given at the 20th Annual Meeting of the Australian Mathematical Society in Perth, May 1976.

In this talk I want to give a survey of some of the areas of research in ordinary differential equations. In view of the vastness of the field, the limited time available and my own inadequacies I will restrict attention to problems in which I have been involved in some way, although I won't be speaking about my own work. Even with this restriction it will be necessary to be somewhat superficial. However, in spite of all these drawbacks, I thought it would assist communication among mathematicians simply to have an idea of what other people are doing. Since this is my purpose I will not mention the quite mild smoothness hypotheses in the statements of some results.

The small oscillations of a mechanical system about a position of equilibrium are described by the solutions of a *linear Hamiltonian system* of differential equations

$$(1) \quad Jx' = Hx,$$

where $x \in R^{2n}$, H is a real $2n \times 2n$ symmetric matrix and

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

The Hamiltonian form is preserved under a linear change of variables $x = Ty$ when the matrix T is *symplectic*, i.e. $T^*JT = J$, where T^* denotes the transpose of T .

The system (1) is said to be *stable* if each solution $x(t)$ is bounded on the whole line $-\infty < t < \infty$. This corresponds to the position of equilibrium being stable. One can characterise stability in a purely algebraic way: the system (1) is stable if and only if there exists a symplectic matrix T such that

$$(2) \quad T^{-1}J^{-1}HT = J^{-1} \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix},$$

where $D = [\lambda_1, \dots, \lambda_n]$ is a real $n \times n$ diagonal matrix. It is desirable to have a somewhat stronger concept of stability which is not destroyed by small perturbations of the coefficient matrix. The system (1) is said to be *strongly stable* if there exists a $\delta > 0$ such that the linear Hamiltonian system

$$Jx' = \tilde{H}x$$

is stable for all real $2n \times 2n$ symmetric matrices \tilde{H} satisfying $|\tilde{H} - H| < \delta$. Strong stability also can be characterised in a purely algebraic way: the system (1) is strongly stable if and only if in the canonical form (2) we have $\lambda_j + \lambda_k \neq 0$ for all j, k . This was shown by Kreĭn (1955, 'if') and Gelfand and Lidskiĭ (1955, 'only if').

Now let $H(\varepsilon)$ be a real symmetric matrix function of ε which is holomorphic at $\varepsilon = 0$, i.e. it can be expanded as a power series with a positive radius of convergence. Suppose also that the unperturbed system $Jx' = H(0)x$ is strongly stable. Then it may be shown that there exists a symplectic matrix function $T(\varepsilon)$, which is holomorphic at $\varepsilon = 0$, such that

$$T^{-1}(\varepsilon)J^{-1}H(\varepsilon)T(\varepsilon) = J^{-1} \begin{pmatrix} D(\varepsilon) & 0 \\ 0 & D(\varepsilon) \end{pmatrix},$$

where $D(\varepsilon) = [\lambda_1(\varepsilon), \dots, \lambda_n(\varepsilon)]$ is a real $n \times n$ diagonal matrix function, which is then also holomorphic at $\varepsilon = 0$. The significance of this result, which is due to Diliberto (1961), is that it provides a justification for formal perturbation procedures. A. Howe (1973) gave a quite different and more general method of proof. It rests on an interesting and powerful method of reducing a matrix function due, independently, to Daleckii and Kato.

Let $P(z)$ be a projection matrix function, i.e. $P^2(z) = P(z)$, which is holomorphic in a simply connected domain G of the complex plane. Then the solution $W(z)$ of the matrix differential equation

$$W' = [P'(z)P(z) - P(z)P'(z)]W$$

such that $W(z_0) = I$, for some $z_0 \in G$, is not only holomorphic and invertible throughout G but also has the property

$$W^{-1}(z)P(z)W(z) = P(z_0) \quad \text{for all } z \in G.$$

For the application to Diliberto's theorem we need only the local version of this theorem, where G is just some neighbourhood of z_0 . However, the global version is no more difficult to prove and is one of the most useful features of the result.

Now let me turn to something quite different, the *method of averaging*. The general idea is to deduce properties of the differential equation

$$(3) \quad x' = \varepsilon f(t, x),$$

where $x \in R^n$ and ε is a small positive parameter, from those of the simpler equation

$$(4) \quad x' = \varepsilon \bar{f}(x),$$

where

$$\bar{f}(x) = \lim_{T \rightarrow \infty} T^{-1} \int_0^T f(t, x) dt$$

is the time mean-value of $f(t, x)$. This idea has long been used in a formal way in celestial mechanics. More recently it has also been applied to non-conservative systems and rigorous results have been obtained. The principal contributor has been Bogoljubov (1934–1958).

Suppose that, for each x , $f(t, x)$ is an *almost periodic* function of t , i.e. it is a uniform limit of trigonometric polynomials $\sum_{k=1}^m (a_k \cos \lambda_k t + b_k \sin \lambda_k t)$. This ensures that the mean value $\bar{f}(x)$ certainly exists. Bogoljubov's three main results are as follows:

(A) If $x(t)$, $\bar{x}(t)$ are solutions of (3), (4) respectively such that $\bar{x}(0) = x(0)$ then, for any $T > 0$,

$$\sup_{0 \leq t \leq T/\varepsilon} |\bar{x}(t) - x(t)| \rightarrow 0 \quad \text{as} \quad \varepsilon \rightarrow +0.$$

(B) If the averaged equation (4) has a constant solution x_0 such that all eigenvalues of the Jacobian matrix $\bar{f}_x(x_0)$ have non-zero real parts then, for all small $\varepsilon > 0$, the equation (3) has in the neighbourhood of x_0 a unique bounded solution $x(t)$. Moreover this solution $x(t)$ is almost periodic and has the same stability properties as the constant solution x_0 of (4).

(C) If the averaged equation (4) has a periodic solution $x_0(\varepsilon t)$ such that $n - 1$ of the characteristic exponents of the corresponding variational equation

$$dy/d\tau = \bar{f}_x[x_0(\tau)]y$$

have non-zero real parts (the remaining characteristic exponent is necessarily zero) then, for all small $\varepsilon > 0$, the equation (3) has in the neighborhood of $x_0(\varepsilon t)$ a unique integral manifold with the same stability properties.

By an *integral manifold* of (3) I mean a manifold $M \subset R^1 \times R^n$ with the property that if $(t_0, x(t_0)) \in M$ for some solution $x(t)$ of (3) and some $t_0 \in R$ then $(t, x(t)) \in M$ for all $t \in R$. Thus it is a manifold made up of complete solutions.

Gihman (1952) showed that the result (A) is a corollary of a much more

general result in which there is no assumption of almost periodicity and no mention of averaging:

Let $x(t)$, $\bar{x}(t)$ be solutions of the differential equations

$$x' = f(t, x), \quad \bar{x}' = f(t, \bar{x}) + g(t, \bar{x}),$$

such that $\bar{x}(0) = x(0)$, and suppose the right sides satisfy the Lipschitz conditions

$$|f(t, x_2) - f(t, x_1)| \leq L |x_2 - x_1|, \quad |g(t, x_2) - g(t, x_1)| \leq L |x_2 - x_1|.$$

If

$$|f(t, x) + g(t, x)| \leq N$$

and if

$$\left| \int_{t_1}^{t_2} g(t, x) dt \right| \leq \gamma \text{ for all } x \text{ and all } t_1, t_2 \text{ with } |t_2 - t_1| \leq 1$$

then

$$|\bar{x}(t) - x(t)| \leq \delta e^{Lt} + \delta(e^{Lt} - 1)/L \text{ for all } t \geq 0,$$

where $\delta = \gamma + (\gamma^2 + 2LN\gamma)^{1/2}$.

When γ is small, δ is also small and $\bar{x}(t)$ is close to $x(t)$ over any prescribed finite interval. This differs from the usual theorems about continuous dependence of solutions on the right side of the differential equation in that we do not require g to be small but only 'integrally small'. I showed (1968) that the result (B) could also be deduced from a more general integrally small perturbation theorem which did not involve averaging. K. J. Palmer (1970) did the same for the result (C), which is a good deal more difficult.

As an example of a rather different type of perturbation problem consider the second order scalar differential equation

$$\epsilon x'' + x' = 0$$

with the boundary conditions

$$x(0) = \alpha, \quad x(1) = \beta.$$

The exact solution is

$$x(t, \epsilon) = [\alpha(e^{-t/\epsilon} - e^{-1/\epsilon}) + \beta(1 - e^{-t/\epsilon})](1 - e^{-1/\epsilon})^{-1}.$$

This shows that $x(t, \epsilon) \rightarrow \beta$ for $0 < t \leq 1$ as $\epsilon \rightarrow +0$. The constant $x_0(t) = \beta$ is a solution of the degenerate equation $x' = 0$ satisfying the boundary condition $x(1) = \beta$. Since the degenerate equation is of lower order than the original equation it is not surprising that one of the boundary conditions should be

lost. That it is the boundary condition at the left endpoint which is lost is due to the fact that the coefficient of x' in the original equation is positive. The limit relation $x(t, \varepsilon) \rightarrow x_0(t)$ as $\varepsilon \rightarrow +0$ holds uniformly in any subinterval $[\delta, 1]$, where $0 < \delta < 1$, and is even uniform outside any neighbourhood of the endpoint $t = 0$ of width $0(\varepsilon^{1/2})$. Such a neighbourhood is often called a *boundary layer*. The name comes from fluid mechanics, where Prandtl first had the ingenious idea of treating the flow of a slightly viscous fluid, such as air or water, as a *singular perturbation* problem. The boundary layer in that case is the region adjacent to the bounding walls in which the velocity of the fluid changes rapidly. One of the reasons for studying singular perturbation problems for ordinary differential equations, although not the only one, is that they serve as models for the more difficult problems associated with partial differential equations.

Of course we require methods which do not depend on the knowledge of exact solutions. Recently K. W. Chang (1976) has treated the boundary value problem

$$(5) \quad \begin{aligned} \varepsilon x'' + A(t, x, \varepsilon)x' &= f(t, x, \varepsilon) \\ x(0) &= \alpha(\varepsilon), \quad x(1) = \beta(\varepsilon), \end{aligned}$$

where $x \in R^n$, A is an $n \times n$ matrix function, f is an n -vector function and ε is again a small positive parameter. He shows that if the degenerate problem

$$(6) \quad \begin{aligned} A(t, x, 0)x' &= f(t, x, 0) \\ x(1) &= \beta(0), \end{aligned}$$

has a solution $x_0(t)$ on the interval $[0, 1]$ and if, for $0 \leq t \leq 1$, every eigenvalue of $A[t, x_0(t), 0]$ has real part $\geq \mu > 0$ then there exist constants $\delta_0 > 0$, $\nu > 0$ such that if $\varepsilon > 0$ is sufficiently small and if $|\alpha(\varepsilon) - x_0(0)| < \delta_0$ then (5) has a solution $x(t)$ on the interval $[0, 1]$ satisfying

$$\begin{aligned} x(t) &= x_0(t) + 0(\varepsilon) + 0(e^{-\nu t/\varepsilon}), \\ x'(t) &= x'_0(t) + 0(\varepsilon) + 0(\varepsilon^{-1} e^{-\nu t/\varepsilon}), \end{aligned}$$

where the boundedness implied by the order symbols is uniform in t .

These results are very close in form to those obtained by Coddington and Levinson (1952) and Wasow (1956) for the case of a scalar equation ($n = 1$), but the method of proof is necessarily quite different. It depends on transforming an associated second order linear differential equation

$$\varepsilon x'' + C(t, \varepsilon)x' + D(t, \varepsilon)x = 0$$

into a block diagonal system

$$\begin{aligned}y' &= P(t, \varepsilon)y \\ \varepsilon z' &= -[C(t, \varepsilon) + \varepsilon P(t, \varepsilon)]z.\end{aligned}$$

This transformation is effected by showing that the solution $P(t, \varepsilon)$ of the matrix Riccati equation

$$\varepsilon P' = -\varepsilon P^2 - C(t, \varepsilon)P - D(t, \varepsilon)$$

which satisfies the initial condition $P(0, \varepsilon) = 0$ is defined and uniformly bounded for $0 \leq t \leq 1$, $0 < \varepsilon \leq \varepsilon_0$.

The general *symmetric matrix Riccati equation*

$$(7) \quad W' = A(t) + WB(t) + B^*(t)W + WC(t)W,$$

where $A^*(t) = A(t)$ and $C^*(t) = C(t)$, plays an important role in control theory and is also closely connected with the theory of linear Hamiltonian systems. If a solution $W(t)$ of (7) is symmetric at one point then it is symmetric throughout its interval of definition. For any two symmetric matrices W_1, W_2 we write $W_1 \geq W_2$ if the difference $W_1 - W_2$ is non-negative definite. This is a partial ordering of the space of symmetric matrices. It is not difficult to show that if $W_1(t), W_2(t)$ are any two symmetric solutions of (7) with $W_1(a) \geq W_2(a)$ then $W_1(t) \geq W_2(t)$ for all t in some two-sided neighbourhood $(a - \delta, a + \delta)$. A. N. Stokes (1974) has proved a remarkable converse to this result.

Consider the matrix differential equation

$$(8) \quad W' = F(t, W),$$

where $F(t, W)$ is an $n \times n$ symmetric matrix for every real t and every $n \times n$ symmetric matrix W . Suppose the equation (8) has the property that if $W_1(t), W_2(t)$ are any two symmetric solutions with $W_1(a) \geq W_2(a)$ then $W_1(t) \geq W_2(t)$ for all t in some two-sided neighbourhood $(a - \delta, a + \delta)$. Then, if $n > 1$, we must have

$$F(t, W) = A(t) + WB(t) + B^*(t)W + WC(t)W,$$

where $A^*(t) = A(t)$ and $C^*(t) = C(t)$.

The restriction $n > 1$ is essential, since *any* scalar differential equation $w' = f(t, w)$ has the order-preserving property.

Let me return now to almost periodic functions. If $f(t)$ is an almost periodic function, is its indefinite integral $F(t) = \int_0^t f(s)ds$ also almost periodic? Since an almost periodic function is bounded, the mean value

$$\lim_{T \rightarrow \infty} T^{-1} \int_0^T f(t) dt$$

must be zero. However this necessary condition is not also sufficient. The root of the difficulty may be seen by considering the Fourier series for $f(t)$:

$$(9) \quad f(t) \sim \sum c_k e^{i\lambda_k t} \quad (\lambda_k \neq 0).$$

If the indefinite integral is almost periodic it has the Fourier series

$$(10) \quad F(t) \sim c_0 + \sum (c_k / i\lambda_k) e^{i\lambda_k t},$$

obtained by term by term integration. To be specific, suppose that $f(t)$ is a quasi-periodic function with two basic frequencies. That is, $f(t) = p(\omega_1 t, \omega_2 t)$, where $p(\varphi_1, \varphi_2)$ is a continuous function with period 2π in each of its variables and ω_1, ω_2 are real numbers whose ratio is irrational, in other words ω_1 and ω_2 are linearly independent over the rationals. This is the simplest type of almost periodic function, which is not periodic. The frequencies λ_k of $f(t)$ all have the form $m_1 \omega_1 + m_2 \omega_2$, where m_1 and m_2 are integers. For suitably chosen m_1, m_2 the frequencies λ_k can be arbitrarily close to zero and the series $\sum |c_k / \lambda_k|^2$ can diverge. Then, by Bessel's inequality, the right side of (10) is not the Fourier series of an almost periodic function. We have in fact a small divisor problem. However Siegel's basic observation about small divisor problems applies here and it may be shown that for almost all pairs (ω_1, ω_2) the denominator $m_1 \omega_1 + m_2 \omega_2$ does not get too close to zero too often. I won't make this statement precise for the present problem, but I will for a related problem.

Consider the quasi-periodic linear differential equation

$$(11) \quad x' = [A + P(\omega t)]x,$$

where $A = [\lambda_1, \dots, \lambda_n]$ is a real diagonal matrix with distinct diagonal elements, and $P(\varphi)$ is an $n \times n$ matrix function of $\varphi = (\varphi_1, \dots, \varphi_h)$ which is real for real φ , has period 2π in each coordinate, and is holomorphic and satisfies the inequality

$$|P(\varphi)| \leq \delta \quad \text{for} \quad |\text{Im } \varphi| < r.$$

Suppose the real vector $\omega = (\omega_1, \dots, \omega_h)$ has the property that for all non-zero vectors $m = (m_1, \dots, m_h)$ with integer coordinates

$$(12) \quad |m_1 \omega_1 + \dots + m_h \omega_h| \geq \gamma |m|^{-\tau},$$

where $\tau > h$ and $\gamma > 0$ are constants.

Then for any $\varepsilon > 0$ there is a corresponding $\delta_0(\varepsilon) > 0$ such that if $\delta \leq \delta_0(\varepsilon)$ the equation (11) has a fundamental matrix of the form

$$X(t) = [I + Q(\omega t)]e^{tB},$$

where $B = [\mu_1, \dots, \mu_n]$ is a real diagonal matrix with $|\mu_k - \lambda_k| \leq \varepsilon$ for all k , and $Q(\varphi)$ is an $n \times n$ matrix function of $\varphi = (\varphi_1, \dots, \varphi_n)$ which is real for real φ , has period 2π in each coordinate, and is holomorphic and satisfies the inequality

$$|Q(\varphi)| \leq \varepsilon \quad \text{for} \quad |\text{Im } \varphi| < r' \leq r.$$

This is a partial generalisation to quasi-periodic linear equations of Floquet's theorem for periodic linear equations. The result was first established by Mitropolskiĭ and Samoilenko (1965) and an improved proof was given by G. C. O'Brien (1976). The proofs use the method of accelerated convergence associated with the names of Kolmogorov, Arnold and Moser. A. Gray (1975) has given an abstract accelerated convergence theorem which applies to many small divisor problems, including this one. But I still hope for a somewhat simpler version.

An approach which is perhaps less general but which I find appealing has been suggested by Hald (1974). Newton's method for solving an operator equation

$$F(x) = 0$$

consists in constructing a sequence of successive approximations

$$x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n),$$

the starting value x_0 being some approximation to the desired root. Under quite general conditions the method is quadratically convergent. In small divisor problems one also needs a rapidly convergent iteration process, to stop the effect of the small divisors from building up. The difficulty is that the derivative may not have a bounded inverse, and in any case the inverse is in general unknown.

Hald suggested solving the pair of equations

$$F(x) = 0$$

$$F'(x)y = I$$

by means of the sequence of successive approximations

$$x_{n+1} = x_n - y_n F(x_n)$$

$$y_{n+1} = y_n - y_n [F'(x_{n+1})y_n - I],$$

where x_0 is an approximation to the desired root and y_0 is a solution, or approximate solution, of $F'(x_0)y = I$. This process is a modification of one due to Moser and, as a result of the modification, is also quadratically convergent.

The small divisor problems I have been discussing are clearly connected with those which Professor Szekeres discussed in his lecture yesterday. I hope that some of you may find a connection between your own work and the other problems which I have mentioned.

References

- N. N. Bogolyubov and Yu. A. Mitropol'skii (1958), *Asymptotic methods in the theory of nonlinear oscillations* (R) (Gosud. Izdat. Fiz.-Mat. Lit., Moscow, 2nd ed. 1958).
- K. W. Chang (1976), 'Singular perturbations of a boundary problem for a vector second order differential equation', *SIAM J. Applied Math.* **30**, 42–54.
- E. A. Coddington and N. Levinson (1952), 'A boundary value problem for a nonlinear differential equation with a small parameter', *Proc. Amer. Math. Soc.* **3**, 73–81.
- W. A. Coppel (1968), 'Dichotomies and reducibility II', *J. Differential Equations* **4**, 386–398.
- S. P. Diliberto (1961), 'On stability of linear mechanical systems', *Proc. Internat. Sympos. Nonlinear Vibrations* (Izdat. Akad. Nauk Ukrain. SSR, Kiev, 1963), Vol. I, pp. 189–203.
- I. M. Gelfand and V. B. Lidskii (1955), 'On the structure of the regions of stability of linear canonical systems of differential equations with periodic coefficients' (R), *Uspehi Mat. Nauk* (N.S.) **10**, no. 1 (63), 3–40.
- I. I. Gihman (1952), 'Concerning a theorem of N. N. Bogolyubov' (R), *Ukrain. Mat. Ž.* **4**, 215–219.
- A. Gray (1975), 'A fixed point theorem for small divisor problems', *J. Differential Equations* **18**, 346–365.
- O. H. Hald (1974/5), 'On a Newton-Moser type method', *Numer. Math.* **23**, 411–426.
- A. Howe (1973), 'Perturbed linear Hamiltonian systems with periodic coefficients', *J. Math. Anal. Appl.* **43**, 31–45.
- M. G. Krein (1955), 'The basic propositions of the theory of λ -zones of stability of a canonical system of linear differential equations with periodic coefficients' (R), *Pamyati Aleksandra Aleksandrovič Andronov* (Izdat. Akad. Nauk SSSR, Moscow, 1955), pp. 413–498.
- Yu. A. Mitropol'skii and A. M. Samoilenko (1965), 'On constructing solutions of linear differential equations with quasiperiodic coefficients by the method of improved convergence' (R), *Ukrain. Mat. Ž.* **17**, no. 6, 42–59.
- G. C. O'Brien (1976), 'Reducibility for real analytic quasi-periodic linear systems', *J. Austral. Math. Soc.* **21**, 289–298.
- K. J. Palmer (1970), 'Averaging and integral manifolds II', *Bull. Austral. Math. Soc.* **2**, 369–399.
- A. N. Stokes (1974), 'A special property of the matrix Riccati equation', *Bull. Austral. Math. Soc.* **10**, 245–253.
- W. R. Wasow (1956), 'Singular perturbations of boundary value problems for nonlinear differential equations of the second order', *Comm. Pure Appl. Math.* **9**, 93–113.

Department of Mathematics,
Institute of Advanced Studies,
Australian National University,
Canberra, ACT.