A Macroscopic Condition for Stability

MORTON M. DENN
University of Delaware, Newark, Delaware

A basic inequality which is necessary and sufficient for macroscopic stability of lumped dynamical systems is applied to obtain further insight into the construction of stability regions, development of stability and control criteria, and meaning and simplified derivation of some conventional results.

Stability analyses for systems described by nonlinear vector ordinary differential equations of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \ \mathbf{f}(\mathbf{0}) = \mathbf{0} \tag{1}$$

where $\mathbf{x}$ is a real vector, generally proceed either by means of a linearized analysis or by application of Liapunov stability theory. In the former case, the stability of the steady state to infinitesmal perturbations is equivalent to the condition that the eigenvalues of the matrix **grad f** (the matrix with components $\partial f_i/\partial x_j$) evaluated at $\mathbf{x} = \mathbf{0}$ have negative real parts. The latter examines stability in macroscopic regions about the origin but requires the construction of a test function for which practical procedures are as yet lacking.

Linearized analyses have two limitations. The most serious from the practical point of view is the restriction to infinitesimal perturbations. The second, largely pedagogical, is the absence of a sound physical interpretation for the mathematical criteria of negative eigenvalues. For example, in the second-order system describing the transient behavior of a stirred tank reactor with a single reaction, the linearized stability conditions are (*1*)

$$\left(\frac{\partial f_1}{\partial x_1}\right)\left(\frac{\partial f_2}{\partial x_2}\right) - \left(\frac{\partial f_1}{\partial x_2}\right)\left(\frac{\partial f_2}{\partial x_1}\right) \geq 0 \tag{2a}$$

$$\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} \leq 0 \tag{2b}$$

The first is derivable from steady state considerations as the requirement that the change of the rate of heat removal with temperature exceed the change of the rate of heat generation (the slope condition). The second, however, has not been given a physical basis.

In the Liapunov approach (*7*), a positive definite function $V(\mathbf{x})$ must be found in a region containing the origin, thus constructing a metric in the phase space. If $\dot{V}$ is negative semidefinite, then phase points can never move further from the origin with respect to that metric than their position at time zero, and the region is stable. A function satisfying both conditions is called a *Liapunov function.* The available methods for construction of Liapunov functions are such that the stability regions are rarely sufficiently large to ensure stability with respect to finite perturbations of a magnitude which might be expected in practice. The application of Liapunov stability theory has recently been reviewed (*6*).

In this paper we reexamine the conditions for finite stability regions from a geometric point of view which is distinct from the Liapunov approach. In this way we are able to gain some further insight into possible methods for construction of stability regions, new stability and control criteria, and the meaning of some of the previously obtained mathematical results.

BASIC INEQUALITY

A region of practical stability (*10*) is defined as a closed region of a phase space from which a phase trajectory can never depart. (Such a region need not contain the origin.) If the closed surface of the region is denoted by $\mathcal{S}$ and the outward normal by $\mathbf{n}$, then it is obvious that a necessary and sufficient condition that a region of phase space be a stability region for a system defined by Equation (1) is

$$\mathbf{n} \cdot \mathbf{f} \leq 0 \text{ everywhere on } \mathcal{S} \tag{3}$$

At surface points where a unique normal is not defined, the inequality must be true in the limit from all directions on the surface.

The usual Liapunov approach is obtained as a special case for which the stability region contains the origin and is bounded by differentiable surfaces defined by constant values of a positive definite function $V(\mathbf{x})$. The normal is then **grad** V, and the basic inequality is equivalent to the negativity of $\dot{V}$. However, all conventional stability results are derivable directly from the basic inequality. For example, if $|\mathbf{x}|$ is small and $\mathbf{f}$ twice differentiable at $\mathbf{x} = \mathbf{0}$, Equation (1) is written as

$$\mathbf{x} = (\mathbf{grad\,f}) \cdot \mathbf{x} + o\,(\mathbf{x}) \tag{4}$$

where **grad f** is evaluated at $\mathbf{x} = \mathbf{0}$ and $\mathbf{o(x)}$ is defined such that

$$\lim_{|\mathbf{x}| \to 0} \frac{|\mathbf{o(x)}|}{|\mathbf{x}|} = 0 \qquad (5)$$

The surface $\mathcal{S}$ is taken as the sphere

$$\mathcal{S}: \mathbf{x} \cdot \mathbf{x} = x^2 \qquad (6)$$

and, because Equation (4) is scale invariant, the norm x is arbitrary. Thus **n** is proportional to **x**, and Equation (3) becomes

$$\mathbf{x} \cdot (\mathbf{grad\,f}) \cdot \mathbf{x} + o(x^2) \leq 0 \qquad (7)$$

Or, normalizing **x** by dividing by x^2, we get

$$\boldsymbol{\xi} \cdot (\mathbf{grad\,f}) \cdot \boldsymbol{\xi} + \frac{o(x^2)}{x^2} \leq 0, \; |\boldsymbol{\xi}| = 1 \qquad (8)$$

By taking the limit as $x \to 0$, the necessary and sufficient condition for stability is

$$\boldsymbol{\xi} \cdot (\mathbf{grad\,f}) \cdot \boldsymbol{\xi} \leq 0, \; |\boldsymbol{\xi}| = 1 \qquad (9)$$

This inequality for all real $\boldsymbol{\xi}$ is equivalent to the requirement that the eigenvalues of **grad f** have negative real parts (see Appendix), a result usually obtained by use of a quadratic Liapunov function (*7*).

The determination of finite regions of stability is equivalent to the construction of surfaces on which the basic inequality is satisfied. For two-dimensional systems, the tracking function method of Paradis and Perlmutter (*10*) is an approach which constructs such a surface as, apparently, is the alternate extreme radius path method of Leathrum et al (*8*). When the surface $\mathcal{S}$ is made up of intersecting hyperplanes, the basic inequality reduces to the choice of parameters for the simultaneous satisfaction of a group of inequalities. If the phase space is sampled to create a test set of points, the parameters defining the stability region can be constructed by using algorithms developed for solving problems of pattern recognition (*9*). This possibility for exploiting the results of extensive research in pattern recognition does not seem to have been recognized previously.

MACROSCOPIC NECESSARY CONDITION

Another macroscopic condition can be derived from the basic inequality by integrating Equation (3) over the surface

$$\int_{\mathcal{S}} (\mathbf{n} \cdot \mathbf{f}) \, d\mathcal{S} \leq 0 \qquad (10)$$

The sum of a sequence of sufficient conditions is not sufficient, so Equation (10) is only necessary. By application of Green's theorem (*2*), denoting the volume of the stability region by $\mathcal{V}$, this is equivalent to

$$\int_{\mathcal{V}} (\mathbf{div\,f}) \, d\mathcal{V} \leq 0 \qquad (11)$$

The divergence of **f**, a scaler, is equal to the sum of the diagonal terms of **grad f**.

A helpful interpretation of this last inequality is obtained by considering a convected volume of phase points in the phase space. At any time t the volume occupied by those points is

$$\mathcal{V}(t) = \int_{\mathcal{V}(t)} d\mathcal{V} \qquad (12)$$

The rate of change of the volume is obtained from the Reynolds transport theorem (*2*) as

$$\frac{d\mathcal{V}}{dt} = \int_{\mathcal{S}(t)} (\mathbf{n} \cdot \mathbf{f}) d\mathcal{S} = \int_{\mathcal{V}(t)} (\mathbf{div\,f}) d\mathcal{V} \qquad (13)$$

so that the macroscopic necessary condition defined by Equation (11) is equivalent to the requirement that the volume in phase space occupied by points on trajectories in a stability region not grow in time. On subsets of the largest possible stability region, the inequality in Equation (11) will be strict, and the volume must shrink, with the exponential rate of decrease governed by the volume average of **div f**. The rate at which macroscopic disturbances are damped out within a stability region is therefore governed by **div f**. That this condition is not sufficient is easily demonstrated by counterexample, for a volume containing the origin can decrease to zero while points within that volume can move arbitrarily far from the origin.

LIMITING CASE

For small volumes about the origin, Equations (11) and (13) become

$$\frac{d\mathcal{V}}{dt} = (\mathbf{div\,f}) + o(\mathcal{V}) \leq 0 \qquad (14)$$

where **div f** is evaluated at $\mathbf{x} = \mathbf{0}$. Within the region where linearization is valid, the shrinking phase volume or damping condition is simply

$$\mathbf{div\,f} \leq 0 \qquad (15)$$

For the two-dimensional case including the chemical reactor, this is the previously unexplained Equation (2*b*). The second stability condition is the dynamic requirement that as the system returns toward the equilibrium temperature (a consequence of the slope condition), the temperature concentration interactions must be such that the disturbance is damped. More precisely, if the particular disturbance is imbedded in a connected family of disturbances, the volume occupied by that family must decrease in time.

The linearized version of the second-order system can always be written in the form

$$\ddot{x} + \mu\dot{x} + kx = 0 \qquad (16)$$

where the necessary and sufficient conditions for stability reduce to $\mu > 0$, $k > 0$. The first of these is the damping condition just developed; the second the slope condition, Equation (2*a*). In the mass-spring-dashpot analogy, the damping term $\mu \geqq 0$ is indeed the viscous damping of the dashpot, while the static slope condition is equivalent to a positive spring constant, also a static quantity. This is the reason that steady state behavior is sufficient to deduce the slope condition for stability, while kinematical considerations are required for damping.

APPLICATION TO CONTROL

Many processes can be described by equations of the form

$$\mathbf{x} = \mathbf{f(x)} + \mathbf{b(x)}\; u \qquad (17)$$

where $u(\mathbf{x})$ is a scaler feedback control variable. For simplicity we take u to be symmetrically bounded:

$$|u| \leq 1 \qquad (18)$$

The fundamental inequality for control stability is then

$$\mathbf{n}\cdot\mathbf{f}+\mathbf{n}\cdot\mathbf{b}\,u\leqq 0 \quad (19)$$

For a given family of surfaces characterized by normals $\mathbf{n}(\mathbf{x})$, the most stable control in the sense of making the inequality in Equation (19) as strong as possible is

$$u(\mathbf{x}) = -\,\mathrm{sgn}\,(\mathbf{n}\cdot\mathbf{b}) \quad (20)$$

where the signum function is defined only for nonzero arguments as

$$\mathrm{sgn}\,(y) = \frac{y}{|y|} = \begin{cases} +1, & y>0 \\ -1, & y<0 \end{cases} \quad (21)$$

Then Equation (19) becomes

$$|\mathbf{n}(\mathbf{x})\cdot\mathbf{b}(\mathbf{x})| \geqq \mathbf{n}(\mathbf{x})\cdot\mathbf{f}(\mathbf{x}) \quad (2)$$

which defines limitations on switching functions for a stable feedback control.

If the uncontrolled system ($u = 0$) is stable, then the inequality in Equation (22) is always satisfied by choosing $\mathbf{n}(\mathbf{x})$ as the normal to stability surfaces for the uncontrolled system, for then $\mathbf{n}(\mathbf{x})\cdot\mathbf{f}(\mathbf{x}) \leqq 0$. If, for example

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\cdot\mathbf{x} \quad (23)$$

where $\mathbf{A}$ is a constant matrix whose eigenvalues have negative real parts, and the surfaces $\mathcal{S}$ are the family of positive definite quadratics

$$\mathbf{x}\cdot\mathbf{Q}\cdot\mathbf{x} = \text{constant} \quad (24)$$

then $\mathbf{n}\cdot\mathbf{f} \leqq 0$ becomes

$$\mathbf{x}\cdot\mathbf{Q}\cdot\mathbf{A}\cdot\mathbf{x}\leqq 0 \quad (25)$$

A family of positive definite surfaces of the type defined in Equation (24) always exists for an arbitrary choice of the symmetric part of a negative definite matrix $\mathbf{Q}\cdot\mathbf{A}$ (7). It has been established previously (3, 4) that controls obtained in this fashion are optimal for a meaningful overall (integral) performance criterion.

OPTIMAL CONTROL

Optimal control problems may often be formulated as

$$\dot{\mathbf{x}} = \mathbf{f}\,(\mathbf{x}, \mathbf{u}) \quad (26)$$

$$\min_{\substack{u(t)\\ 0\leqq t\leqq \theta}} \int_0^\theta F(\mathbf{x}, \mathbf{u})\,dt \quad (27)$$

Here $\mathbf{u}$ is a vector of control functions. A function $S(\mathbf{x})$ may be defined as

$$S(\mathbf{x}) = \min_{\substack{u(\tau)\\ t\leqq\tau\leqq\theta}} \int_{t(\mathbf{x})}^\theta F(\mathbf{x}, \mathbf{u})\,dt \quad (28)$$

that is, the remaining contribution to the objective function along an optimal path evaluated at point $\mathbf{x}$. In minimum time control, for example, $S(\mathbf{x})$ would be the minimum time from $\mathbf{x}$ to the origin, and constant values of S would define isochrones or surfaces of constant minimum time to the origin. Then the dynamic programming formulation (4, 5) leads to the Hamilton-Jacobi-Bellman equation

$$\min_{\mathbf{u}(t)} \; [F(\mathbf{x}, \mathbf{u}) + (\mathbf{grad}\,S)\cdot\mathbf{f}(\mathbf{x}, \mathbf{u})] = 0 \quad (29)$$

We denote **grad** S by $\mathbf{n}$. Within the region where the system is controllable to the origin from arbitrary initial states, the constant values of $S(\mathbf{x})$ will form closed surfaces, and, since F is positive definite, $\mathbf{n}\cdot\mathbf{f} \leqq 0$. Furthermore, if F is independent of $\mathbf{u}$, then the Hamilton-Jacobi-Bellman equation requires that the optimal control be the most stable in the sense of making $\mathbf{n}\cdot\mathbf{f}$ as negative as possible, which means that the motion is in the direction of steepest descent of S consistent with the system equations. If F does depend on $\mathbf{u}$, then the optimal control is not the most stable but represents a compromise between maximum stability relative to the cost surface and instantaneous cost of control.

LITERATURE CITED

1. Aris, Rutherford, "Introduction to the Analysis of Chemical Reactors," Prentice Hall, Englewood Cliffs, N. J. (1965).
2. ———, "Vectors, Tensors, and the Basic Equations of Fluid Mechanics," Prentice Hall, Englewood Cliffs, N. J. (1962).
3. Denn, M. M., *AIChE J.*, **13**, 926 (1967).
4. ———, "Optimization by Variational Methods," McGraw-Hill, New York (1969).
5. Dreyfus, S. E., "Dynamic Programming and the Calculus of Variations," Academic Press, New York (1965).
6. Gurel, O., and Leon Lapidus, *Ind. Eng. Chem.*, **60**, No. 6, 12 (1968).
7. LaSalle, J. P., and S. Lefschetz, "Stability by Liapunov's Direct Method with Applications," Academic Press, New York (1961).
8. Leathrum, J. F., E. F. Johnson, and Leon Lapidus, *AIChE J.*, **10**, 16 (1964).
9. Nilsson, N. J., "Learning Machines," McGraw-Hill, New York (1965).
10. Paradis, W. O., and D. D. Perlmutter, *AIChE J.*, **12**, 131 (1966).

Manuscript received August 14, 1968; revision received January 6, 1969; paper accepted January 8, 1969.

APPENDIX

It is shown that the inequality

$$\mathbf{x}\cdot\mathbf{A}\cdot\mathbf{x}\leqq 0 \quad (30)$$

for arbitrary $\mathbf{x}$ implies that the real parts of all eigenvalues of $\mathbf{A}$ are negative. The eigenvalue equation is

$$\mathbf{A}\cdot\mathbf{y} = \lambda\mathbf{y} \quad (31)$$

where $\mathbf{y}$ is an eigenvector and λ the coresponding eigenvalue. In general, it is necessary to break λ and $\mathbf{y}$ into real and imaginary parts:

$$\lambda = \lambda_R + i\,\lambda_I \quad (32a)$$

$$\mathbf{y} = \mathbf{y}_R + i\,\mathbf{y}_I \quad (32b)$$

Thus, Equation (31) is written as

$$\mathbf{A}\cdot(\mathbf{y}_R + i\mathbf{y}_I) = (\lambda_R + i\lambda_I)(\mathbf{y}_R + i\mathbf{y}_I) \quad (33)$$

and by equating real and imaginary parts, we get

$$\mathbf{A}\cdot\mathbf{y}_R = \lambda_R\,\mathbf{y}_R - \lambda_I\mathbf{y}_I \quad (34a)$$

$$\mathbf{A}\cdot\mathbf{y}_I = \lambda_R\,\mathbf{y}_I + \lambda_I\,\mathbf{y}_R \quad (34b)$$

The inner product of Equation (34*a*) with $\mathbf{y}_R$ and Equation (34*b*) with $\mathbf{y}_I$ yields

$$\mathbf{y}_R\cdot\mathbf{A}\cdot\mathbf{y}_R = \lambda_R\,|\mathbf{y}_R|^2 - \lambda_I\mathbf{y}_R\cdot\mathbf{y}_I\leqq 0 \quad (35a)$$

$$\mathbf{y}_I\cdot\mathbf{A}\cdot\mathbf{y}_I = \lambda_R\,|\mathbf{y}_I|^2 + \lambda_I\mathbf{y}_R\cdot\mathbf{y}_I\leqq 0 \quad (35b)$$

and, by adding Equations (35*a*) and (35*b*)

$$\lambda_R(|\mathbf{y}_R|^2 + |\mathbf{y}_I|^2)\leqq 0 \quad (36)$$

or

$$\lambda_R\leqq 0 \quad (37)$$