Stability and asymptotic estimates in nonautonomous linear differential systems

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

STABILITY AND ASYMPTOTIC ESTIMATES IN NONAUTONOMOUS LINEAR DIFFERENTIAL SYSTEMS*

GUSTAF SÖDERLIND[†] AND ROBERT M. M. MATTHEIJ[‡]

Abstract. A new theory is presented, in which a generalized kinematic similarity transformation is used to diagonalize linear differential systems. No matrices of Jordan form are needed. The relation to Lyapunov's classical stability theory is explored, and asymptotic estimates of fundamental solutions are given. Finally, some possible numerical applications of the presented theory are suggested.

1. Introduction. In this paper, we consider linear systems of ordinary differential equations

$$\dot{x} = A(t)x + g(t), \qquad x \in \mathbb{R}^m \tag{1.1}$$

to model the propagation of perturbations in a general nonlinear system

$$\dot{y} = f(t, y), \tag{1.2}$$

where $y \in \mathbb{R}^m$ and $f: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$. If $z \in \mathbb{R}^m$ satisfies the perturbed equation

$$\dot{z} = f(t, z) + g(t), \tag{1.2'}$$

we note that the difference $x = z - y$ satisfies (1.1), where $A(t)$ is the "average Jacobian"

$$A(t) = \int_0^1 J(t, y + \theta x) \, d\theta.$$

Here the $m \times m$ matrix $J(\cdot, \cdot)$ is the partial derivative of f with respect to its second argument. Although a linearization is not necessary in order to establish (1.1), the matrix $A(t)$ depends, by construction, not only on t but also upon x and y. This limits the validity of (1.1) as a model for the error propagation in (1.2), since $A(t)$ may not be uniformly bounded with respect to x. However, with the additional requirement that f satisfies the Lipschitz condition

$$\|f(t, z) - f(t, y)\| \le L \|z - y\| \quad \forall t, y, z$$

one easily shows that

$$\|A(t)\| \le L \quad \forall t. \tag{1.3}$$

We note that the Lipschitz condition can be relaxed; it is sufficient that the condition holds over a convex domain $D \subset \mathbb{R}^m$, i.e. whenever $y, z \in D$.

Under mild conditions, the homogeneous problem $\dot{x} = A(t)x$ has a continuously differentiable fundamental solution matrix Φ, i.e.

$$\dot{\Phi} = A(t)\Phi. \tag{1.4}$$

Using this operator, the solution of (1.1) can, in terms of some given initial condition $x(0)$, be written

$$(1.5) \qquad x(t) = \Phi(t)\Phi^{-1}(0)x(0) + \Phi(t)\int_0^t \Phi^{-1}(\tau)g(\tau)\,d\tau.$$

We remark that if Φ is a fundamental matrix over the semi-infinite interval $[0, \infty)$, then Φ^{-1} exists on any finite subinterval of $[0, \infty)$.

The object of the paper is to estimate the solution x given by (1.5). In particular, we are interested in asymptotic estimates and stability, i.e. we want to find estimates of $\|x(t)\|$ as well as of $\|\Phi(t)\Phi^{-1}(0)\|$. We will derive these estimates for a monotonic but otherwise unspecified norm. In particular cases we will consider the Hölder norms.

The estimates for global error propagation that we obtain are similar to corresponding results derived by using the logarithmic norm, [6], [8] and [20]. Although the latter estimates are sharp for "short-range" error propagation, our estimates are generally better for large t. Thus, they can be viewed as a complement to the traditional logarithmic norm bounds on the error.

In §2, basic concepts will be introduced and classical results reviewed. In §3 we consider various choices of a fundamental solution. The fundamental solution will then be decomposed into a normalized direction matrix and a size matrix which satisfies a differential equation kinematically similar to (1.1) [11]. We also prove a new diagonalization theorem, demonstrating that any matrix can be brought to diagonal form using a (time-dependent) transformation of Lyapunov type. This result is of fundamental importance since it allows a unified treatment of all linear systems, whether A be constant, defective or time-dependent. It is particularly useful in the latter case, when a Jordan form no longer has a clear meaning. It should be noted that the techniques presented here are of equal importance to initial value problems and boundary value problems.

In §4 we derive the asymptotic error estimates for IVP's by considering the Lyapunov transformation and its adjoint equation. Finally, in §5 we consider some applications of the presented theory.

2. Differential inequalities and logarithmic norms. "Classical" estimates of the solution to (1.1) are obtained from the differential inequality[1]

$$\frac{d}{dt}\|x\| \le \|A(t)\|\|x\| + \|g\|.$$

Due to the Lipschitz constant $\|A\|$ being positive, these estimates are in practice useless, since they fail to provide information about the actual growth or decay rate in (1.1). The situation was greatly improved by the introduction of logarithmic norms [8], [6], [20]. In terms of the logarithmic norm of the matrix A, defined by

$$(2.1) \qquad \mu[A] = \lim_{h \to 0+} \frac{\|I + hA\| - 1}{h},$$

solutions to (1.1) can be estimated from

$$(2.2) \qquad \frac{d}{dt}\|x\| \le \mu[A(t)]\|x\| + \|g\|.$$

More precisely, we can state the following lemma [20].

---

[1]Since $\|x\|$ may be only piecewise differentiable, the derivative of $\|x\|$ is to be interpreted as a right-hand derivative.

LEMMA 1. *Let $x(t)$ be a solution of $\dot{x} = A(t)x + g(t)$. Then $\|x(t)\| \leq \xi(t)$, where the scalar function ξ satisfies the differential equation*

$$(2.2') \qquad \dot{\xi} = \mu[A(t)]\xi + \|g(t)\|$$

with the initial condition $\xi(0) = \|x(0)\|$.

While the Lipschitz constant is always positive, the logarithmic norm may be negative. This implies that sufficient conditions for classical stability notions can easily be expressed in terms of $\mu[A]$, see e.g. [6]. Instead of going into details, we shall only summarize some useful basic properties of the logarithmic norm that can be found elsewhere in the literature (see [8] and [20]).

We define the spectral abscissa of a matrix A by

$$(2.3) \qquad \alpha[A] = \max_i \text{Re}(\lambda_i)$$

where $\lambda_1, \cdots, \lambda_m$ are the eigenvalues of A.

LEMMA 2. *Let A and B be square matrices. Let γ be a nonnegative real number and z be a complex number. Then*
  a) $\alpha[A] \leq \mu[A]$;
  b) $\mu[\gamma A] = \gamma \mu[A]$, $\gamma \geq 0$;
  c) $\mu[A + zI] = \mu[A] + \text{Re}(z)$;
  d) $-\|A\| \leq \mu[A] \leq \|A\|$;
  e) $\mu[A + B] \leq \mu[A] + \mu[B]$.
Furthermore, if Λ is a diagonal matrix and the norm $\|\cdot\|$ is monotonic, [1], then
  f) $\mu[\Lambda] = \alpha[\Lambda]$.

LEMMA 3. *Let A be a constant quadratic matrix. Then*
  a) $\|e^{At}\| \leq e^{\mu[A]t}$;
  b) $\mu[A] = \lim_{h \to 0+} \log\|e^{Ah}\|/h$.

If A depends on t, we can derive nonautonomous counterparts to the statements in Lemma 3:

LEMMA 3'. *Let Φ be a continuously differentiable fundamental solution satisfying* (1.4). *Then*
  a) $\|\Phi(t)\Phi^{-1}(\tau)\| \leq \exp\int_\tau^t \mu[A(s)]\,ds$;
  b) $\mu[A(t)] = \lim_{h \to 0+} \log\|\Phi(t+h)\Phi^{-1}(t)\|/h$.

Proof. Part a) follows immediately from (1.5) and Lemma 1. In part b), note that $\Phi(t+h) = \Phi(t) + h\dot{\Phi}(t) + o(h) = (I + hA(t))\Phi(t) + o(h)$. Hence $\Phi(t+h)\Phi^{-1}(t) = I + hA(t) + o(h)$, and, as a consequence of (2.1),

$$(2.4) \qquad \lim_{h \to 0+} \frac{\|\Phi(t+h)\Phi^{-1}(t)\| - 1}{h} = \mu[A(t)].$$

We also obtain

$$(2.5) \qquad \frac{d}{dt}\|\Phi(t)\Phi^{-1}(\tau)\|\Big|_{t=\tau} = \mu[A(\tau)].$$

It follows that $\|\Phi(t+h)\Phi^{-1}(t)\| = 1 + h\mu[A(t)] + o(h)$ as $h \to 0+$. The result then follows by taking logarithms and letting $h \to 0+$. $\square$

The significance of Lemma 3b) and 3'b) is that error bounds obtained by using Lemma 1 are sharp (with respect to the particular choice of norm) for short term error propagation. However, over long intervals, the logarithmic norm may sometimes give

gross overestimates. We illustrate these matters by considering the nonautonomous homogeneous equation

(2.6)
$$\dot{x} = A(t)x,$$
$$\|x(0)\| = 1 \qquad \text{(unit initial error)}.$$

By Lemma 1,

$$\|x(h)\| \le \xi(h) = \exp \int_0^h \mu[A(s)]\, ds.$$

Since $x(h) = \Phi(h)\Phi^{-1}(0)x(0)$, $\xi(h)$ must clearly majorize $\|\Phi(h)\Phi^{-1}(0)\|$, which is the largest possible value of $\|x(h)\|$ given that $\|x(0)\| = 1$. By (2.5), the Maclaurin expansions of $\xi(h)$ and $\|\Phi(h)\Phi^{-1}(0)\|$ agree to first-order terms in h. The minimal margin is therefore

$$\xi(h) - \|\Phi(h)\Phi^{-1}(0)\| = o(h)$$

as $h \to 0+$, showing that Lemma 1 indeed yields sharp results for short-term error propagation.

Estimates of the asymptotic behavior based on the logarithmic norm give useful results only if the vector norm has been chosen with extreme care. Consider, for example, the constant coefficient system

(2.7)
$$\dot{x} = \begin{bmatrix} -1 & 10 \\ 0 & -2 \end{bmatrix} x$$

with the matrix exponential

(2.8)
$$e^{At} = \begin{bmatrix} e^{-t} & 10(e^{-t} - e^{-2t}) \\ 0 & e^{-2t} \end{bmatrix}.$$

It is immediately clear that for any choice of norm and initial condition, the asymptotic behavior of the solution is $\|x(t)\| \sim e^{-t}$. Yet, if we choose the maximum norm, we find that $\mu_\infty[A] = 9$. Thus Lemma 1 yields $\xi(t) = e^{9t}$, whereas

(2.8′)
$$\|e^{At}\|_\infty = 11e^{-t} - 10e^{-2t}.$$

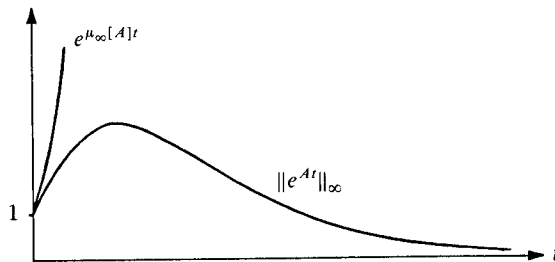These bounds are illustrated in Fig. 1.

FIG. 1.

In a constant coefficient system it is a fairly straightforward task to construct a norm giving useful asymptotic estimates. Thus if $T^{-1}AT$ is diagonal, we can define

$\|x\|_T = \|T^{-1}x\|_\infty$, from which we can derive $\mu_T[A] = \alpha[A]$. We now generalize this technique to defective and time-dependent systems by using a local nonsingular coordinate transformation

$$(2.9) \qquad x(t) = T(t) y(t).$$

We want to estimate the solution in terms of a given *monotonic* norm $\|\cdot\|_p$, the *global norm* of the solution. We define a *local* (time-dependent) *norm* $\|\cdot\|_T$ by

$$(2.10) \qquad \|x(t)\|_T = \|T^{-1}(t)x(t)\|_p = \|y(t)\|_p.$$

We first estimate $\|x\|_T$, then these results are transformed back to estimates with respect to the fixed global norm. The following inequalities are readily established:

$$(2.11) \qquad \begin{aligned} &\|x\|_p \le \|T\|_p \|y\|_p = \|T\|_p \|x\|_T, \\ &\|x\|_T \le \|T^{-1}\|_p \|x\|_p. \end{aligned}$$

From the definition of the logarithmic norm (2.1) it follows that

$$(2.12) \qquad \mu_T[A] = \mu_p[T^{-1}AT].$$

When (2.9) is applied to (1.1), we obtain the differential equation

$$(2.13) \qquad \dot{y} = (T^{-1}AT - T^{-1}\dot{T})y + T^{-1}g,$$

from which we derive the differential inequality

$$(2.14) \qquad \frac{d}{dt}\|y\|_p \le \mu_p[T^{-1}AT - T^{-1}\dot{T}]\|y\|_p + \|T^{-1}g\|_p.$$

While (2.2) still remains valid for the fixed time-independent global norm, (2.14) clearly shows that for the local norm (cf. (2.12)),

$$(2.14') \qquad \frac{d}{dt}\|x\|_T \le \mu_T[A - \dot{T}T^{-1}]\|x\|_T + \|g\|_T.$$

Note the term $\dot{T}T^{-1}$, which accounts for the time-dependence of the local norm. Estimates of $\|x\|_T$ can now be obtained by applying Lemma 1 to (2.13), and then transformed back to estimates of $\|x\|_p$ by means of (2.11). We shall see that we can choose the coordinate transformation (2.9) in such a way that $\mu_T[A - \dot{T}T^{-1}]$ is significantly smaller than $\mu_p[A]$, thereby permitting estimates with better asymptotic properties. The price to be paid for this advantage is that we lose sharpness for short-term error propagation.

Before concluding this section, we point out that the following inequality,

$$(2.15) \qquad \frac{d}{dt}\|T\|_p \le \mu_p[T^{-1}\dot{T}]\|T\|_p = \mu_p[\dot{T}T^{-1}]\|T\|_p,$$

which follows directly from the identities $\dot{T} = TT^{-1}\dot{T} = \dot{T}T^{-1}T$, is sometimes useful in deriving the asymptotic estimates.

3. Kinematic eigenvalues and the Lyapunov transformation.
Throughout the paper we shall assume that the matrix function $A(t)$ satisfies the following assumptions:

Assumption A1. $A(t)$ is uniformly bounded with respect to t, i.e. $\|A(t)\|_p \le L, \forall t$.

Assumption A2. There exists a continuously differentiable fundamental solution matrix Φ satisfying $\dot{\Phi} = A\Phi, \forall t$.

It follows from A1 that no solution to the homogeneous problem $\dot{x} = Ax$ can grow faster than $\exp(Lt)$. Similarly, no homogeneous solution can decay faster than $\exp(-Lt)$. In order to measure the asymptotic behavior of solutions, it is therefore convenient to use the concept of characteristic exponents or type numbers, [4, p. 50], [11] and [17, p. 165].

DEFINITION. The *generalized Euclidean characteristic exponent* of a vector function $f(t)$ is defined by

$$(3.1) \qquad \chi(f) = \overline{\lim_{t \to \infty}} \frac{\log\|f(t)\|_2}{t}.$$

If f and g are vector functions and γ is a scalar function of t, then it is clear from the definition that the generalized characteristic exponent satisfies the following rules:

$$(3.2) \qquad \chi(f+g) \leq \max(\chi(f), \chi(g))$$

with equality if $\chi(f) \neq \chi(g)$, and

$$(3.3) \qquad \chi(\gamma f) \leq \chi(\gamma) + \chi(f).$$

DEFINITION. Let $\chi(f) = \chi(g) = \chi_0$. If for some nonzero constants γ_1 and γ_2 we have

$$\chi(\gamma_1 f + \gamma_2 g) < \chi_0,$$

we call f and g *exponentially linearly dependent*. Otherwise they are exponentially linearly independent.

Example. Consider the constant coefficient system

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x,$$

with fundamental solutions

$$\Phi = \begin{bmatrix} e^t & e^{-t} \\ e^t & -e^{-t} \end{bmatrix}, \qquad \Psi = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}.$$

Ψ serves well as a fundamental solution when t is small, but away from the origin its columns rapidly become almost linearly dependent. In fact, they are exponentially linearly dependent. The columns of Φ, on the other hand, are orthogonal for all t. We remark that given a fundamental matrix Φ, one can construct a new fundamental solution by postmultiplying Φ by any constant nonsingular matrix M. It should also be noted that since Φ^{-1} exists for all t, the columns of Φ are linearly independent in such a way that every time-dependent linear combination $\Phi(t)\gamma(t) \neq 0$ if $\gamma(t) \neq 0$ for all t. This "spatial" linear independence is much stronger than the "functional" linear independence $\phi(t)\gamma \not\equiv 0$ for every constant vector γ.

DEFINITION. Let $\Phi = (\phi_1, \phi_2, \cdots, \phi_m)$, where $\phi_j \colon \mathbb{R} \to \mathbb{R}^m$, be a fundamental matrix. Φ is said to be *normal* in the sense of Lyapunov [17, p. 169], [4, p. 52] if

$$\sum_{j=1}^{m} \chi(\phi_j) \text{ is minimal.}$$

Clearly, in the previous example, Φ is normal with $\Sigma\chi(\phi_j) = 0$, whereas Ψ, with $\Sigma\chi(\psi_j) = 2$ is not normal.

In the following analysis we shall be concerned with the following particular choice of Φ:

Assumption A3. Φ is normal, and its columns have been permuted so that Φ can be partitioned columnwise into

$$(3.4) \qquad \Phi = \left[\Phi_1, \Phi_2, \cdots, \Phi_q \right], \qquad 1 \le q \le m,$$

where for every j, each column in the submatrix Φ_j has the same generalized characteristic exponent χ_j. The characteristic exponents are assumed to be arranged in descending order, i.e. $\chi_j > \chi_{j+1}$.

It can be shown that a normal fundamental matrix always exists, and so Assumption A3 is always satisfied for some fundamental matrix. Unless otherwise stated, in the sequel we will deal exclusively with fundamental matrices satisfying A3.

Remark. Observe that if Φ satisfies A3, then so does $\tilde{\Phi} = \Phi M$, where M is a nonsingular block lower triangular matrix partitioned conformally with (3.4), i.e.

$$(3.5) \qquad M = \begin{bmatrix} M_{11} & 0 & \cdots & 0 \\ M_{21} & M_{22} & & \\ \vdots & & \ddots & \vdots \\ & & & 0 \\ M_{q1} & & \cdots & M_{qq} \end{bmatrix}, \qquad \det M_{jj} \ne 0.$$

Any result induced by A3 therefore remains qualitatively, but not necessarily quantitatively, the same for Φ and $\tilde{\Phi}$.

PROPOSITION 4. *Let Φ satisfy Assumption* A3. *Then the columns within each submatrix Φ_j are exponentially linearly independent.*

Proof. Suppose there is a nonzero vector γ for which $\chi(\Phi_j\gamma) < \chi_j$. Then we can construct a new fundamental matrix $\tilde{\Phi}$ by replacing an arbitrary column in Φ_j by $\Phi_j\gamma$. We then have

$$\sum_{j=1}^{m} x(\tilde{\phi}_j) < \sum_{j=1}^{m} \chi(\phi_j),$$

thus contradicting the assumption that Φ is normal. $\qquad \square$

We shall now decompose Φ into a *direction matrix* T and a diagonal *size matrix* D,

$$(3.6) \qquad \Phi = TD$$

where $T = (t_1, t_2, \cdots, t_m)$ has columns of unit Euclidean norm,

$$(3.7) \qquad t_j^T t_j = 1.$$

Then, since $\phi_j = t_j d_{jj}$, we have

$$(3.8) \qquad d_{jj} = \|\phi_j\|_2.$$

The following properties of T and D immediately follow from Assumption A2 and the differentiability of the Euclidean norm:

PROPOSITION 5. *T and D are nonsingular and continuously differentiable.*

Now, since $\dot{\Phi} = A\Phi$, we obtain $\dot{T}D + T\dot{D} = ATD$, or

$$\dot{T} = AT - T\dot{D}D^{-1}.$$

Denote the diagonal matrix $\dot{D}D^{-1}$ by Λ. Then

$$(3.9) \qquad \dot{T} = AT - T\Lambda,$$

$$(3.10) \qquad \dot{D} = \Lambda D.$$

Thus we have proved the following diagonalization theorem:

THEOREM 6. *For every matrix function $A(t)$ satisfying Assumptions* A1 *and* A2 *there exists a continuously differentiable nonsingular matrix T such that $\Lambda = T^{-1}AT - T^{-1}\dot{T}$ is diagonal. Moreover, under Assumption* A3, *a possible choice of T is given by* (3.6)–(3.8).

COROLLARY 7. *For every matrix function $A(t)$ satisfying Assumptions* A1 *and* A2 *there exists a continuously differentiable nonsingular matrix T such that the differential equation*

$$(3.11) \qquad \dot{x} = A(t)x + g(t)$$

is decoupled by the coordinate transformation $x = Ty$ into a system of scalar differential equations

$$(3.12) \qquad \dot{y} = \Lambda(t)y + T^{-1}(t)g(t).$$

The fundamental solutions Φ and D, associated with (3.11) *and* (3.12), *respectively, are related by the same transformation, i.e.,* $\Phi = TD$.

A coordinate transformation $x = Ty$ is called a Lyapunov transformation, [10, p. 117], under the conditions that

   (i) T is uniformly bounded,
   (ii) $\dot{T}$ is continuous and uniformly bounded,
   (iii) T^{-1} is uniformly bounded.

We shall see that the direction matrix T obtained by the decomposition (3.6) satisfies conditions (i) and (ii) but not always (iii). Thus it is well-known that a defective matrix cannot be transformed to diagonal form with a transformation T satisfying (iii) over a semi-infinite interval. However, there are only two reasons for considering condition (iii). Firstly, T may not become singular for any finite t. By Proposition 5 this cannot occur in our case. Secondly, if any uniform upper bound of T^{-1} appears in some estimate of the solution, it clearly has to be finite. However, such a bound will not be needed in our estimates. Thus (iii) is unnecessarily restrictive, and unless otherwise stated, we shall replace that condition by the weaker requirement

   (iii') T^{-1} exists on every finite interval,

which, according to Proposition 5 is always satisfied. We call the resulting transformation, satisfying (i), (ii) and (iii'), a *generalized Lyapunov transformation*. In particular, we shall refer to (3.9) as a *diagonalizing Lyapunov transformation* (boundedness of $\dot{T}$ will be established in Propositions 9 and 10). The systems (3.11) and (3.12) are said to be *kinematically similar* (cf. [11] and [4, p. 54]), and we call the diagonal elements $\lambda_i(t)$ of $\Lambda(t)$ the *kinematic eigenvalues of A with respect to T*. Let

$$(3.13) \qquad S^T = T^{-1}.$$

Then S and T provide the left and right kinematic eigenvalues of A. We remark that T and Λ are *not unique* unless we specify exactly which normal fundamental matrix Φ is to be used for the construction of T. Asymptotic properties, however, are uniquely determined as we shall see in Propositions 11 and 12.

We now illustrate our results by considering the defective system

$$(3.14) \qquad \dot{x} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} x, \qquad t \geq 0,$$

which has a fundamental matrix

$$\Phi = \begin{bmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{bmatrix}.$$

It is easily verified that

$$(3.14') \qquad T = \begin{bmatrix} 1 & \dfrac{t}{\sqrt{1+t^2}} \\ 0 & \dfrac{1}{\sqrt{1+t^2}} \end{bmatrix}, \qquad S^T = \begin{bmatrix} 1 & -t \\ 0 & \sqrt{1+t^2} \end{bmatrix},$$

corresponding to

$$(3.14'') \qquad \Lambda = \begin{bmatrix} -1 & 0 \\ 0 & -1 + \dfrac{t}{1+t^2} \end{bmatrix}, \qquad D = \begin{bmatrix} e^{-t} & 0 \\ 0 & \sqrt{1+t^2}\, e^{-t} \end{bmatrix}.$$

Note that the kinematic eigenvalues are not constant despite the fact that the original system has constant coefficients. This is a consequence of the diagonalizing Lyapunov transformation being time-dependent. Also note that the fundamental matrices Φ and D both exhibit asymptotic growths e^{-t} and te^{-t}.

It is clearly seen that S^T is not uniformly bounded with respect to t. We expressly state that this is not a deficiency of the presented theory. It merely reflects the fact that for *any* choice of fundamental matrix Φ, the space spanned by its columns collapses as $t \to \infty$. This property is inherited by the kinematic eigensystem T which is aligned with the directions of the linearly independent solutions ϕ_j. Finally, we point out that there are nonautonomous systems with distinct eigenvalues that behave in a similar way. Thus, for instance, the system

$$\dot{x} = \begin{bmatrix} -1 & 1 \\ 0 & -1 + \dfrac{1}{1+t} \end{bmatrix} x, \qquad t \geq 0$$

has a fundamental matrix

$$\Phi = \begin{bmatrix} e^{-t} & -\dfrac{t^2}{2} e^{-t} \\ 0 & (1+t)e^{-t} \end{bmatrix}.$$

It is clear that for any choice of Φ, S^T will be $O(t)$ as $t \to \infty$. This "pseudodefective" behavior is due to the two eigenvalues of A approaching a defective pair as $t \to \infty$.

The kinematic eigenvalues and eigenvectors have a number of interesting properties:

PROPOSITION 8. *The kinematic eigenvalue λ_j is equal to the Rayleigh quotient formed by the corresponding kinematic eigenvector t_j and A, i.e.*

$$(3.15) \qquad \lambda_j = t_j^T A t_j.$$

Proof. Differentiating (3.7), we find that $t_j^T \dot{t}_j = 0$. By (3.9), $\dot{t}_j = A t_j - t_j \lambda_j$. Thus, $0 = t_j^T A t_j - \lambda_j$.

PROPOSITION 9. *All kinematic eigenvalues satisfy the inequalities*

$$(3.16) \qquad -L \leq -\mu_2[-A] \leq \lambda_j \leq \mu_2[A] \leq L.$$

Proof. For all x with $x^T x = 1$ we have that $-\mu_2[-A] \leq x^T A x \leq \mu_2[A]$. The inequalities then follow from (3.15) and part d) of Lemma 2. □

PROPOSITION 10. $\dot{T}$ *is uniformly bounded with respect to t.*

Proof. The result immediately follows from equation (3.9) and the boundedness of A, T and Λ. □

PROPOSITION 11. *The characteristic exponents are preserved by the diagonalizing Lyapunov transformation, i.e. if* $\Phi = TD$, *then* $\chi(\phi_j) = \chi(d_j)$.
Proof.

$$\chi(\phi_j) = \overline{\lim_{t \to \infty}} \frac{1}{t} \log \|\phi_j\|_2 = \overline{\lim_{t \to \infty}} \frac{1}{t} \log d_{jj} = \chi(d_j). \qquad \square$$

PROPOSITION 12. $\chi(\phi_j)$ *can be expressed as the "infinite average"*

$$(3.17) \qquad \chi(\phi_j) = \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \lambda_j(s) \, ds.$$

Proof. By (3.10), $\dot{d}_j = \lambda_j d_j$. Integrating yields

$$d_j(t) = \exp \int_0^t \lambda_j(s) \, ds \, d_j(0).$$

Hence Proposition 11 gives

$$\chi(d_j) = \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \lambda_j(s) \, ds = \chi(\phi_j). \qquad \square$$

DEFINITION. The *kinematic spectral abscissa* of A with respect to T is defined by

$$(3.18) \qquad \alpha_T[A] = \max_j \lambda_j.$$

We then have

PROPOSITION 13. *Let* $\|\cdot\|_p$ *be monotonic and let* $\|\cdot\|_T$ *be defined by (2.10) where T is a diagonalizing Lyapunov transformation. Then*

$$\alpha_T[A] = \mu_T[A - \dot{T}T^{-1}] = \mu_p[\Lambda].$$

Proof. From (2.12) we obtain $\mu_T[A - \dot{T}T^{-1}] = \mu_p[T^{-1}AT - T^{-1}\dot{T}] = \mu_p[\Lambda]$ by (3.9). Since $\|\cdot\|_p$ is monotonic, Lemma 2f) gives $\mu_p[\Lambda] = \alpha_T[A]$. □

It is clear that $\alpha_T[A]$ has strong implications as to the stability of (1.1). Not only is the kinematic spectral abscissa closely related to the characteristic exponents, but it appears explicitly in (2.14) and (2.14'). Thus we have uniform stability if $\alpha_T[A] \leq 0$ and uniform asymptotic stability if $\alpha_T[A] \leq -\alpha < 0$ for all t. We note that these results cannot be concluded from corresponding conditions for the spectral abscissa $\alpha[A]$ if the system is nonautonomous. These questions will be further discussed in §4.

An interesting consequence of Theorem 6 is

THEOREM 14 (exponential representation theorem). *Every fundamental solution admits the exponential representation*

$$(3.19) \qquad \Phi(t)\Phi^{-1}(\tau) = T(t)\exp\int_\tau^t \Lambda(s)\,ds S^T(\tau)$$

whenever Assumptions A1 *and* A2 *are satisfied.*

Proof. Take M so that ΦM satisfies A3. Then

$$\Phi(t)\Phi^{-1}(\tau) = \Phi(t)MM^{-1}\Phi^{-1}(\tau) = T(t)D(t)D^{-1}(\tau)S^T(\tau).$$

Since $\dot{D} = \Lambda D$, (3.19) follows from

$$(3.20) \qquad D(t) = \exp\int_\tau^t \Lambda(s)\,ds D(\tau),$$

and the representation

$$(3.21) \qquad \Phi(t) = T(t)D(t). \qquad\qquad \square$$

Remark. Note that (3.19) is a generalization to the nonautonomous case of the corresponding formula in the diagonalizable constant coefficient case. Thus, if there exists a static similarity transformation that takes A to diagonal form,

$$0 = AT - T\Lambda$$

where $T^{-1} = S^T$ and $A = T\Lambda S^T$, then

$$e^{A(t-\tau)} = Te^{\Lambda(t-\tau)}S^T.$$

It is clearly seen that this formula appears as a special case in Theorem 14. In the nonautonomous case, however, it is well known that

$$\Phi(t)\Phi^{-1}(\tau) = \exp\int_\tau^t A(s)\,ds$$

if and only if A commutes with its derivative, i.e. when $\dot{A}A - A\dot{A} = 0$. The importance of Theorem 14 is that we indeed still have an exponential representation, even if the commutativity condition is not satisfied. It should be noted that this is made possible by the kinematic similarity transformation to diagonal form, and the Lyapunov-type relation $\Phi = TD$, where the fundamental solution D associated with the decoupled system (3.12) always has an exponential representation (3.20). We finally remark that the kinematic diagonalization is a transformation of global character; the case when A is defective locally requires no special attention and no matrices of Jordan form are needed.

We shall now turn to the question of how the matrix $T^{-1} = S^T$ behaves for increasing t. We have already seen that globally defective or pseudo-defective systems will (in general) cause an $O(t^\beta)$ growth for some power $\beta > 0$, due to the inherent structure of the problem. In Theorem 14, however, we would like to avoid any exponential growth of S^T in (3.19), or any exponential linear dependence in the columns of T, so that the exponential behavior is due to Λ only.

Introduce the notation $\phi = \det\Phi$, $\tau = \det T$, $\sigma = \det S^T$ and $\delta = \det D$.

LEMMA 15 [4. p. 53]. $\Sigma\chi(\phi_j) \geq \chi(\phi) \geq -\chi(\phi^{-1})$.

Proof. Since $1 = \phi\phi^{-1}$, we have $0 \leq \chi(\phi) + \chi(\phi^{-1})$ from which the last inequality follows. For the first inequality, note that $\phi = \tau\delta \Rightarrow \chi(\phi) \leq \chi(\tau) + \chi(\delta)$. However, the

normalization of the columns of T gives $|\tau| \leq 1 \Rightarrow \chi(\tau) \leq 0$. Since $\delta = \Pi \|\phi_j\|_2$, we find that $\chi(\delta) \leq \Sigma \chi(\phi_j)$, thus completing the proof. □

In order to show that S^T does not grow exponentially, we have to show that $\chi(\sigma) \leq 0$. Since $\sigma\tau = 1$, we obtain

$$(3.22) \qquad 0 \leq \chi(\tau) + \chi(\sigma).$$

However, $\chi(\tau) \leq 0$, and it follows that S^T does not grow exponentially if and only if $\chi(\tau) = \chi(\sigma) = 0$.

DEFINITION. Under Assumption A1, the system $\dot{x} = Ax$ is said to be *regular* if there exists at least one fundamental solution Φ satisfying

$$(3.23) \qquad \Sigma \chi(\phi_j) = -\chi(\phi^{-1}).$$

It is clear that a fundamental solution satisfying (3.23) must be normal, and, without loss of generality, we may assume that it has the form described in Assumption A3. We now have

THEOREM 16. *Let the system $\dot{x} = Ax$ be regular and let T be a diagonalizing Lyapunov transformation with inverse S^T. Let $\tau = \det T$ and $\sigma = \det S^T$. Then $\chi(\tau) = \chi(\sigma) = 0$. In other words: the columns of T are exponentially linearly independent and S^T does not grow exponentially as t increases.*

Proof. Note that $\sigma = \phi^{-1}\delta = \phi^{-1}\Pi\|\phi_j\|_2$. Hence

$$(3.24) \qquad \chi(\sigma) \leq \chi(\phi^{-1}) + \Sigma \chi(\phi_j).$$

Since the system is regular, (3.23) gives $\chi(\sigma) \leq 0$. (3.22) together with $\chi(\tau) \leq 0$ then yields $\chi(\tau) = \chi(\sigma) = 0$. □

PROPOSITION 17. *The system $\dot{x} = Ax$ is regular only if*

$$(3.25) \qquad \lim_{t \to \infty} \frac{1}{t} \int_0^t \operatorname{tr} A(s)\,ds$$

exists.

Proof. It is well known [5, p. 67] that ϕ satisfies the differential equation $\dot{\phi} = \operatorname{tr} A(t)\phi$. Hence

$$\chi(\phi) = \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \operatorname{tr} A(s)\,ds.$$

Similarly, $\psi = \phi^{-1}$ satisfies the adjoint equation $\dot{\psi} = -\operatorname{tr} A^T(t)\psi$, from which we derive

$$\chi(\phi^{-1}) = \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t -\operatorname{tr} A(s)\,ds = -\underline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \operatorname{tr} A(s)\,ds.$$

Since a regular system has $\chi(\phi) = -\chi(\phi^{-1})$, the existence of the limit (3.25) follows. □

PROPOSITION 17'. *The system $\dot{y} = \Lambda y$ is regular if and only if*

$$(3.26) \qquad \lim_{t \to \infty} \frac{1}{t} \int_0^t \operatorname{tr} \Lambda(s)\,ds$$

exists.

Proof. The "if" part follows from the decoupled structure of the system $\dot{y} = \Lambda y$. It is clearly seen that a scalar system $\dot{d} = \lambda d$ is regular if and only if $\chi(d) = -\chi(d^{-1})$. □

THEOREM 18. *If $\dot{x} = Ax$ is regular, then so is the transformed system $\dot{y} = \Lambda y$, and*

$$(3.27) \qquad \lim_{t \to \infty} \frac{1}{t} \int_0^t \operatorname{tr} A(s) - \operatorname{tr} \Lambda(s) \, ds = 0.$$

Proof. From $\delta^{-1} = \phi^{-1}\tau$ and Theorem 16 it follows that $\chi(\delta^{-1}) \leq \chi(\phi^{-1})$. Hence, if A is regular,

$$\chi(\delta) \leq \sum \chi(\phi_j) = \chi(\phi) = -\chi(\phi^{-1}) \leq -\chi(\delta^{-1}).$$

By Lemma 15 we must have $\chi(\delta) \geq -\chi(\delta^{-1})$, and so the regularity of the transformed system follows. Thus we have $\chi(\phi) = \chi(\delta)$, where (cf. Proposition 12) $\chi(\phi)$ and $\chi(\delta)$ are given by the limits (3.25) and (3.26) respectively. Alternatively, (3.27) may be derived from the two adjoint differential equations

$$\dot{\tau} = (\operatorname{tr} A - \operatorname{tr} \Lambda)\tau, \qquad \dot{\sigma} = (\operatorname{tr} \Lambda - \operatorname{tr} A)\sigma$$

together with $\chi(\tau) = \chi(\sigma) = 0$. $\qquad \square$

Remark. Note that $\operatorname{tr} A$ is equal to the sum of the eigenvalues of A. Thus (3.27) states that the kinematic eigenvalues of A are "close" to the eigenvalues in the infinite average.

If A is permitted to grow exponentially, it is simple to construct problems where S^T grows exponentially, see e.g. [7, p. 12]. However, under Assumption A1, we have found that regularity is a sufficient condition for S^T to grow at most at a polynomial rate. Necessary conditions are still an open question, and at present we are not aware of any system where S^T does grow exponentially. Indeed, in Lyapunov's classical example of an irregular system, [4, pp. 53–54], we actually have a uniformly bounded S^T. One should note, however, that the class of regular systems is very wide. Thus, for instance, all systems with constant or periodic coefficients fall in this class. Irregular systems have fundamental solutions containing elements with a quite odd behavior, e.g. like $\exp(t \sin \log t)$.

4. Asymptotic estimates and condition numbers. We shall derive asymptotic estimates by applying the theory of §3 to the differential inequalities in §2. We begin by giving an estimate for $\|\Phi(t)\Phi^{-1}(0)\|_p$, i.e. we consider the homogeneous problem (2.6).

LEMMA 19. *Let $\|\cdot\|_p$ and $\|\cdot\|_q$ be dual Hölder norms (i.e. $1/p + 1/q = 1$) and assume that A is a rank one matrix, $A = uv^T$. Then*

$$(4.1) \qquad \|A\|_p = \|u\|_p \|v\|_q.$$

Proof.

$$\|A\|_p = \sup_{\|x\|_p = 1} \|Ax\|_p = \sup_{\|x\|_p = 1} \|uv^Tx\|_p = \|u\|_p \sup_{\|x\|_p = 1} |v^Tx|.$$

By Hölder's inequality, $|v^Tx| \leq \|v\|_q \|x\|_p$ with equality for some x. Hence $\|A\|_p = \|u\|_p \|v\|_q$. $\qquad \square$

THEOREM 20. *Let $\|\cdot\|_p$ and $\|\cdot\|_q$ be dual Hölder norms. In addition to Assumptions A1 and A2, assume that $\chi(\phi_1) > \chi(\phi_j)$ for $j \geq 2$. Then, as $t \to \infty$,*

$$(4.2) \qquad \|\Phi(t)\Phi^{-1}(0)\|_p \approx \|t_1(t)\|_p \|s_1^T(0)\|_q \exp \int_0^t \lambda_1(s) \, ds,$$

where t_1 and s_1^T are the first column and row, respectively, of the matrices T and S^T.

Proof. From the exponential representation in Theorem 14, we see that $\Phi(t)\Phi^{-1}(0)$ can be written as a sum of rank one matrices,

$$(4.3) \qquad \Phi(t)\Phi^{-1}(0) = \sum_{j=1}^{m} t_j(t) s_j^T(0) \exp \int_0^t \lambda_j(s)\, ds.$$

If $\chi(\phi_1) > \chi(\phi_j)$ for $j \geq 2$, then by Proposition 12, the terms 2 through m will be exponentially small compared to the first term of the sum in the right-hand side of (4.3). Hence, for large t,

$$\Phi(t)\Phi^{-1}(0) \approx t_1(t) s_1^T(0) \exp \int_0^t \lambda_1(s)\, ds,$$

and the result follows by application of Lemma 19. $\qquad \square$

Remark. The most important application of Theorem 20 is to systems satisfying

$$(4.4) \qquad \lambda_1(t) > \lambda_j(t), \qquad t \geq 0, \quad 2 \leq j \leq m.$$

We then have $\chi(\phi_1) > \chi(\phi_j)$ for $j \geq 2$ if and only if (4.4) holds uniformly with respect to t. It is worth noting, however, that Theorem 20 and its proof remain valid for systems satisfying the weaker requirement

$$(4.5) \qquad \lim_{t \to \infty} \int_0^t \lambda_1(s) - \lambda_j(s)\, ds = +\infty, \qquad 2 \leq j \leq m,$$

although the dominated terms may no longer be exponentially small. Thus (4.4) does not have to hold uniformly in t, and Theorem 20 can also be applied in the defective case. Also note that if (4.4)–(4.5) hold, then the asymptotic behavior is determined (sharply) by the kinematic spectral abscissa, $\alpha_T[A] = \lambda_1$. Finally, note that s_1^T is only evaluated at $t = 0$ in the estimate (4.2), showing that a uniform upper bound of S^T is not needed.

We now illustrate Theorem 20 by returning to the problem (2.7). A can be brought to diagonal form by a static similarity transformation $0 = AT - T\Lambda$, with

$$\lambda_1 = -1, \quad t_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad s_1^T = (1 \quad 10).$$

Thus (4.2) yields the asymptotic estimate

$$\|e^{At}\|_\infty \approx \|t_1\|_\infty \|s_1^T\|_1 e^{\lambda_1 t} = 11 e^{-t}$$

in agreement with (2.8′). For the defective system (3.14), we obtain kinematic eigenvalues (3.14″) that (after permutation) satisfy (4.5). The kinematic spectral abscissa is $-1 + t/(1+t^2)$, corresponding to the left kinematic eigenvector $(0 \ \sqrt{1+t^2})$ appearing in the second row of S^T in (3.14′). Evaluating this vector at $t = 0$, we obtain $(0 \ 1)$. Hence, for the Euclidean norm we have

$$\|e^{At}\|_2 \approx 1 \cdot 1 \cdot \sqrt{1+t^2}\, e^{-t} = \sqrt{1+t^2}\, e^{-t} \approx t e^{-t},$$

a result which is asymptotically sharp for large t.

Next, we turn to estimates of $\|x\|_p$. If $x = Ty$ is a diagonalizing Lyapunov transformation, then by Corollary 7 we have

$$(4.6) \qquad \dot{y} = \Lambda y + S^T g.$$

$\|y\|_p$ can be estimated from (2.14) or by application of Lemma 1, i.e. $\|y\|_p \leq \eta$, where

$$(4.7) \qquad \dot{\eta} = \alpha_T[A]\eta + \|S^T g\|_p$$

with the initial condition $\eta(0) = \|y(0)\|_p$. Integration of (4.7) yields

$$(4.8) \qquad \|y(t)\|_p \leq \|y(0)\|_p \exp \int_0^t \alpha_T[A(s)]\, ds$$

$$+ \int_0^t \left\{ \exp \int_\tau^t \alpha_T[A(s)]ds \right\} \|S^T(\tau) g(\tau)\|_p d\tau.$$

Using (2.11) to transform this estimate into an estimate for $\|x(t)\|_p$, we get

$$(4.9) \qquad \|x(t)\|_p \leq \|T(t)\|_p \|S^T(0) x(0)\|_p \exp \int_0^t \alpha_T[A(s)]\, ds$$

$$+ \|T(t)\|_p \int_0^t \left\{ \exp \int_\tau^t \alpha_T[A(s)]\, ds \right\} \|S^T(\tau) g(\tau)\|_p d\tau$$

whereas direct application of Lemma 1 to (1.1) gives

$$(4.9') \quad \|x(t)\|_p \leq \|x(0)\|_p \exp \int_0^t \mu_p[A(s)]\, ds + \int_0^t \left\{ \exp \int_\tau^t \mu_p[A(s)]\, ds \right\} \|g(\tau)\|_p d\tau.$$

Note that in (4.9), $S^T(\tau)g(\tau)$ is the kinematic spectral projection of $g(\tau)$ onto the local coordinate system at time τ, having the columns of $T(\tau)$ as basis vectors. By applying (4.9) to the homogeneous case, we readily establish the following (usually cruder) alternative to the result of Theorem 20,

$$(4.10) \qquad \left\| \Phi(t)\Phi^{-1}(0) \right\|_p \leq \|T(t)\|_p \|S^T(0)\|_p \exp \int_0^t \alpha_T[A(s)]\, ds$$

whereas (4.9') or Lemma 3'a) yields

$$(4.10') \qquad \left\| \Phi(t)\Phi^{-1}(0) \right\|_p \leq \exp \int_0^t \mu_p[A(s)]\, ds.$$

Note that the bound (4.10) holds without the special assumptions of Theorem 20 or restrictions like (4.4)–(4.5). Although for a given norm it is usually superior to (4.10') for asymptotic purposes, it should be clear that (4.10) does not necessarily give the optimal exponential behavior unless some restrictions of the mentioned type are imposed. Thus,

$$\chi\left(\left\| \Phi(t)\Phi^{-1}(0) \right\|_p \right) \sim \max_i \chi(\phi_i) = \max_i \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \lambda_i(s)\, ds$$

$$\leq \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \max_i \lambda_i(s)\, ds = \overline{\lim_{t \to \infty}} \frac{1}{t} \int_0^t \alpha_T[A(s)]\, ds.$$

The advantage, however, is that $\alpha_T[A]$ is, in principle, a computable quantity; the characteristic exponents, on the other hand, are in practice virtually impossible to compute.

As for the stability properties of the homogeneous problem, we can state the following theorems. We leave the first theorem without proof since the uniform boundedness of T implies that the bounds (4.10) and (4.10') have the same generic

structure, and the corresponding results are well known in the case of the logarithmic norm, [6, p. 59].

THEOREM 21. *Let $\alpha_T[A]$ be the kinematic spectral abscissa of A with respect to the diagonalizing Lyapunov transformation T. Then the zero solution of $\dot{x} = A(t)x$ is*

   a) *stable if $\overline{\lim}_{t\to\infty} \int_0^t \alpha_T[A(s)]\,ds < \infty$;*

   b) *asymptotically stable if $\overline{\lim}_{t\to\infty} \int_0^t \alpha_T[A(s)]\,ds = -\infty$;*

   c) *uniformly stable if $\alpha_T[A(t)] \leq 0$ for $t \geq 0$;*

   d) *uniformly asymptotically stable if $\alpha_T[A(t)] \leq -\alpha < 0$ for $t \geq 0$.*

THEOREM 22. *Let T be a diagonalizing Lyapunov transformation, and assume that $\alpha_T[A] \leq 0$ for all $t \geq 0$. Then the quadratic form $x^T(TT^T)^{-1}x$ is a Lyapunov function for the system $\dot{x} = A(t)x$ provided that T^{-1} is uniformly bounded for $t \geq 0$.*

Proof. Note that $x = Ty$ gives $\dot{y} = \Lambda(t)y$. Define

$$V(y) = y^T y = \|y\|_2^2.$$

By assumption, $\alpha_T[A] = \mu_2[\Lambda] \leq 0$, implying that $V(y)$ is a Lyapunov function for the y-system, i.e. $\dot{V} \leq 0$. Transforming back, $y = T^{-1}x$ now gives

$$V(y) = x^T T^{-T} T^{-1} x = x^T (TT^T)^{-1} x,$$

and the theorem is proved.   □

Remark. A time-dependent function $V(t,x)$ is a Lyapunov function if $\dot{V} \leq 0$ along the solution under consideration, *and* if there are positive definite time-invariant functions $U(x)$, $W(x)$ such that $U(x) \leq V(t,x) \leq W(x)$. Therefore, we have to require that T^{-1} be uniformly bounded in *this* application.

Quantitatively, (4.10) is superior to (4.10') for t large enough to make

$$\|T(t)\|_p \|S^T(0)\|_p \exp \int_0^t \alpha_T[A(s)] - \mu_p[A(s)]\,ds \leq 1$$

or, equivalently, when

(4.11) $$\log\|T(t)\|_p\|S^T(0)\|_p \leq \int_0^t \mu_p[A(s)] - \alpha_T[A(s)]\,ds.$$

In the constant coefficient diagonalizable case, (4.11) reduces to

$$\log \kappa_p[T] \leq \left(\mu_p[A] - \alpha[A]\right)t$$

where $\kappa_p[T]$ is the condition number of the eigenvector matrix t with respect to $\|\cdot\|_p$. In §2 we saw that the logarithmic norm gives sharp estimates initially, but in this case (4.10) is preferable for $t \geq t^*$, where

$$t^* = \frac{\log \kappa_p[T]}{\mu_p[A] - \alpha[A]}.$$

The quantity in the denominator is called the logarithmic inefficiency of the norm $\|\cdot\|_p$, [20]. In the general case we may, because of the normalization of the columns of T, think of $\|s_j^T\|_2$ as a condition number of the corresponding column t_j in T. It follows that $\|S^T\|_p$ is an indication of the local conditioning of the Lyapunov transformation. *We therefore suggest that the matrix (3.5) be taken to minimize $\kappa_p[T]$ in a suitable way.*

The significance of this is clearly seen in the problem

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 - \dfrac{1}{t+1} \end{bmatrix} x, \qquad t \ge 0,$$

which has fundamental solutions

$$\Phi = \begin{bmatrix} e^t & 0 \\ 0 & \dfrac{1}{t+1} e^t \end{bmatrix}, \qquad \Psi = \begin{bmatrix} e^t & e^t \\ 0 & \dfrac{1}{t+1} e^t \end{bmatrix}.$$

Although both matrices satisfy Assumption A3, the first one gives $T = S^T = I$, whereas the latter choice yields a matrix S^T which grows like $O(t)$ as $t \to \infty$.

It is possible to derive differential inequalities where the condition number does appear explicitly. Indeed, in a closely related context, albeit with somewhat different aims, it has been proposed by Dahlquist (private communication) to consider the quantity $\xi = \|T\|_p \|y\|_p$ directly. Thus if $x = Ty$, then $\|x\|_p \le \xi$. Upon differentiation of ξ, one obtains

$$\dot{\xi} = \|y\|_p \frac{d}{dt} \|T\|_p + \|T\|_p \frac{d}{dt} \|y\|_p.$$

The derivatives are, as usual, interpreted as right-hand derivatives. Using (2.15) for the first term and (2.14) for the second, we find

(4.12) $$\dot{\xi} \le \left(\mu_p[T^{-1}AT - T^{-1}\dot{T}] + \mu_p[T^{-1}\dot{T}] \right) \xi + \kappa_p[T] \|g\|_p$$

with the initial condition $\xi(0) = \kappa_p[T(0)] \|x(0)\|_p$. Thus, if T is a diagonalizing Lyapunov transformation,

(4.12′) $$\dot{\xi} \le \left(\alpha_T[A] + \mu_p[T^{-1}\dot{T}] \right) \xi + \kappa_p[T] \|g\|_p.$$

In general, the term $\mu_p[T^{-1}\dot{T}]$ will prevent us from obtaining estimates with the desired exponential behavior, and so (4.12) is best suited for transformations other than the diagonalizing Lyapunov transformation considered in this paper. Comparing (4.12′) and (4.7), we see that in both cases, a small $\kappa_p[T]$ is needed in order to avoid a too large amplification of the forcing term $g(t)$ when it is projected onto the columns of T. In the case of a diagonalizing Lyapunov transformation suitably chosen to minimize $\kappa_p[T]$, a large or growing $\kappa_p[T]$ (such as in the defective case) merely reflects an inherent "ill-conditioning" of the differential system that is inevitable. We repeat that this is not a consequence of our transformation technique; the equations (3.11) and (3.12) appearing in Corollary 7 are completely equivalent. Thus nothing can be gained by forcing S^T to be uniformly bounded at the price of transforming the system to a rather artificial time-dependent Jordan form. Instead we suggest that one interpret $\kappa_p[T]$ as a condition number indicating how well one can distinguish different homogeneous solutions asymptotically as $t \to \infty$. Formally, one may impose conditions that would define T and Λ uniquely, but at present it is not clear what additional properties the "best possible" diagonalizing Lyapunov transformation should possess. Knowing that any transformation of this type does give the optimal exponential behavior in terms of the generalized characteristic exponents, we leave this question open.

5. Applications. In this final section, we shall hint at some possible areas of application of the presented theory. First, we shall briefly discuss how the kinematic spectral abscissa can be computed.

The practical computation of $\alpha_T[A]$ is based on the matrix differential equation (3.9). Note that the structure of (3.9) admits a convenient incorporation of a shift. Thus, if

$$(5.1) \qquad \tilde{A} = A + \beta I, \qquad \tilde{\Lambda} = \Lambda + \beta I,$$

then

$$(5.2) \qquad \dot{T} = \tilde{A}T - T\tilde{\Lambda}.$$

In order to compute $\alpha_T[A]$, we have to compute the maximum kinematic eigenvalue and the corresponding kinematic eigenvector (λ_1 and t_1 say). By (5.2) these quantities satisfy

$$(5.3) \qquad \dot{t}_1 = (\tilde{A} - \tilde{\lambda}_1 I) t_1$$

together with the normalization requirement (3.7), i.e.

$$(5.4) \qquad t_1^T t_1 = 1.$$

An approximation to $\lambda_1 = \tilde{\lambda}_1 - \beta$ is obtained by discretizing (5.3), for instance by the backward Euler method, in which case one gets

$$(5.5) \qquad t_{1,n+1} - t_{1,n} = h(\tilde{A}_{n+1} - \tilde{\lambda}_{1,n+1} I) t_{1,n+1}.$$

Here h is the time-step, and the subscript n indicates an approximation at time t_n. Next, we rearrange (5.5) to obtain

$$h\tilde{\lambda}_{1,n+1} t_{1,n+1} = -(I - h\tilde{A}_{n+1}) t_{1,n+1} + t_{1,n}.$$

This formula is the basis for the iteration

$$(5.6) \qquad h\tilde{\lambda}_{1,n+1}^{k+1} t_{1,n+1}^{k+1} = -(I - h\tilde{A}_{n+1}) t_{1,n+1}^k + t_{1,n}.$$

In each iteration, $\tilde{\lambda}_{1,n+1}^{k+1}$ and $t_{1,n+1}^{k+1}$ are defined by imposing (5.4). Under mild assumptions, the iteration converges with an appropriate shift β. Note that (5.6) is essentially a power iteration, but with an inhomogeneous term taking the time-dependence of the kinematic eigenvector into account. Several other discretizations and iterations are also possible. This is currently being investigated and will be reported elsewhere. Finally, we point out that the mentioned technique is good only for the computation of approximations to the dominant kinematic eigenpair, i.e., we cannot compute the whole T matrix this way. However, this is not the purpose of our analysis. Moreover, if a full transformation were to be computed, one might expect some numerical difficulties. As an alternative, one may consider the possibility of using orthogonal transformation matrices. Thus, if

$$(5.7) \qquad \Phi = QR$$

is the QR factorization of a fundamental matrix Φ, it is easily verified that

$$(5.8) \qquad \dot{Q} = AQ - QU, \qquad \dot{R} = UR,$$

where R and U are upper triangular matrices. (5.8) can then be regarded as a kinematic Schur form of the matrix A, but note that the diagonal elements of U are, in general,

not equal to our kinematic eigenvalues. We remark that the existence of the transformation (5.8) is a classical result, cf. [4, p. 54].

Because of the strong relation between $\alpha_T[A]$ and the stability of the system, one of the most important numerical applications of the presented theory is to monitor the mathematical stability of the problem when it is solved numerically. It is well known, [9] and [14], that frequently used numerical methods for solving stiff initial value problems may sometimes produce erroneous results. This is a consequence of the difference between mathematical stability on the one hand and numerical stability on the other hand. Thus most methods for stiff problems are numerically stable also in large portions of the right half-plane. As a result, the numerical method sometimes follows a (mathematicially) unstable particular solution without ever detecting this instability. However, by numerically computing $\alpha_T[A]$ along the approximate particular solution, such instabilities are easily detected. We therefore propose that $\alpha_T[A]$ be computed so as to implement a stability check in stiff codes that would increase their reliability. Also note that once $\alpha_T[A]$ is computed, it is simple to find an approximation to the integral

$$\int_0^t \alpha_T[A(s)]\,ds$$

which gives information about the global stability properties, cf. Theorem 21.

A second possible application is to estimate the global truncation error in the numerical integration. Instead of solving a variational equation of the type (3.11), we consider a kinematically similar system (3.12). If $\alpha_T[A]=\lambda_1$, then

$$(5.9) \qquad\qquad \dot{y}_1 = \alpha_T[A]\,y_1 + s_1^T g$$

defines the asymptotically dominant component in the y-system. Since

$$x = \sum_{j=1}^m t_j y_j,$$

we see that $t_1 y_1$ is the global error component in the direction with the least damping in the x-system. This component can be estimated from the scalar equation

$$(5.10) \qquad\qquad \dot{\eta} = \alpha_T[A]\eta + \left|s_1^T g\right|.$$

If the global error components in the other directions can be neglected, then

$$\|x\|_p \sim \eta\|t_1\|_p$$

and, in particular, $\|x\|_2 \sim \eta$. It is clear, however, that η is neither a bound nor an estimate of the global error, but rather a "global error indicator" which is not very robust. In some preliminary computations we have, nevertheless, obtained some fairly reasonable results by solving

$$(5.10') \qquad\qquad \dot{\eta} = \alpha_T[A]\eta + \left\|s_1^T\right\|_q \|g\|_p$$

instead of (5.10). The norms are dual Hölder norms, and (5.10') corresponds to projecting the local error vector g entirely onto the t_1 direction. While (5.10') still does not give more than an indication of the global error, it is more robust and only involves quantities associated with λ_1 (s_1 satisfies the adjoint of (5.3)). To obtain a global error

bound, we need $\max_i \|s_i^T\|_q$ or $\|S^T\|_p$, and at present we have not been able to compute these quantities (cf. (4.7)).

We would finally like to give an example showing that the presented theory is also useful for problems not satisfying Assumption A1, i.e. when the matrix A is no longer uniformly bounded with respect to t or other parameters. The application of our technique to such problems can be justified, although the theory is considerably more complicated. Thus the limits defining the characteristic exponents may not exist, and it is also questionable whether one may consider a linear equation (1.1) as a model for the error propagation in a nonlinear equation (1.2).

We consider the following very simple turning point problem

$$(5.11) \qquad\qquad \varepsilon u'' + 2u' + 0 \cdot u = 0$$

over the interval $(-\infty, \infty)$. We are especially interested in how well it is possible to distinguish the two linearly independent solutions of (5.11). In particular, we would like to compute the condition number of the transformation matrix T as $t \to -\infty, +\infty$ and at the turning point $t = 0$. To this end, we rewrite (5.11) as the first-order system

$$(5.12) \qquad\qquad \begin{bmatrix} u' \\ \sqrt{\pi\varepsilon}\, u'' \end{bmatrix} = \begin{bmatrix} 0 & 1/\sqrt{\pi\varepsilon} \\ 0 & -2t/\varepsilon \end{bmatrix} \begin{bmatrix} u \\ \sqrt{\pi\varepsilon}\, u' \end{bmatrix}.$$

Introduce

$$(5.13) \qquad\qquad E(t) = e^{-t^2/\varepsilon}, \qquad I(t) = \frac{1}{\sqrt{\pi\varepsilon}} \int_{-\infty}^{t} E(\tau)\, d\tau.$$

The factor $\sqrt{\pi\varepsilon}$ appears in (5.13) to normalize I so that $I(\infty) = 1$. It is now easily verified that (5.12) has a fundamental solution matrix

$$(5.14) \qquad\qquad \Phi = \begin{bmatrix} 1 & I \\ 0 & E \end{bmatrix}.$$

Since $E \to 0$ as $t \to \infty$, we see that the T matrix associated with (5.14) has a rapidly growing inverse S^T. We therefore try to improve this behavior by considering another fundamental matrix ΦM. Indeed, for

$$M = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

we obtain

$$(5.14') \qquad\qquad \Phi M = \begin{bmatrix} 1-I & I \\ -E & E \end{bmatrix},$$

a fundamental solution having much better properties. Thus, when (5.14') is decomposed into its direction matrix T and size matrix D, we get

$$(5.15) \qquad\qquad T = \begin{bmatrix} \dfrac{1-I}{\sqrt{(1-I)^2 + E^2}} & \dfrac{I}{\sqrt{I^2 + E^2}} \\[4mm] \dfrac{-E}{\sqrt{(1-I)^2 + E^2}} & \dfrac{E}{\sqrt{I^2 + E^2}} \end{bmatrix}$$

and

(5.16)
$$D = \begin{bmatrix} \sqrt{(1-I)^2 + E^2} & 0 \\ 0 & \sqrt{I^2 + E^2} \end{bmatrix}.$$

After some limit calculations, one obtains Table 1, valid uniformly with respect to ε.

TABLE 1

t	$-\infty$	0	∞
T	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\dfrac{1}{\sqrt{5}}\begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
S^T	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\dfrac{\sqrt{5}}{4}\begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
$\kappa_\infty[T]$	1	3	1

The size functions d_{11} and d_{22} in (5.16) are illustrated in Fig. 2.
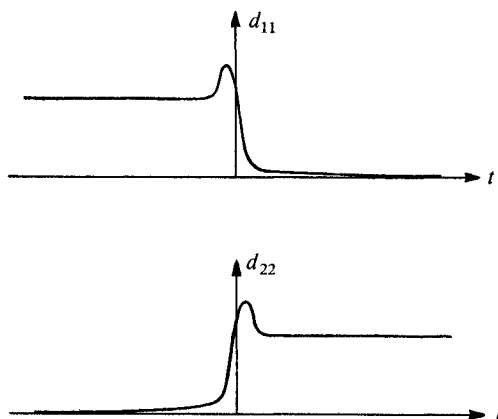

FIG. 2

In Fig. 2 we see that our turning point problem has a dichotomy, [7], in a very general sense. From the table it is clear that the direction matrix T is well behaved. At the turning point this local coordinate system rapidly flips an angle of $\pi/2$ from one orthogonal system to another. This change takes place quicker as $\varepsilon \to 0$, and consequently $\dot{T}$ is not uniformly bounded with respect to ε. The kinematic eigenvalues with respect to T are $0(\varepsilon^{-1/2})$ in a neighborhood of the turning point $t = 0$.

We remark that the special scaling with respect to ε used to derive the first-order system (5.12) is necessary to obtain these good results (an alternative would be to choose a linear combination M that depends in a nonuniform way on ε). Our interpretation of this is that a proper minimization of $\kappa[T]$ implies a proper scaling with respect to the perturbation parameter.

We shall now demonstrate the importance of directional well-conditioning for boundary value problems, and apply some results from [15] to this turning point problem. Consider the BVP

(5.17) $\dot{x} = Ax + g,$ $a \le t \le b,$

(5.18) $M_a x(a) + M_b x(b) = c,$

where c is a vector and M_a, M_b are square matrices which are normalized such that $\max(\|M_a\|_p, \|M_b\|_p) = 1$. It was shown in [15] that the sensitivity of the solution x with respect to the boundary condition (5.18) can be quantified by the following *condition number* ($\|\cdot\|_p$ is a Hölder norm)

$$(5.19) \qquad\qquad CN_p := \max\|\Phi(t)Q^{-1}\|_p,$$

where

$$(5.20) \qquad\qquad Q = M_a\Phi(a) + M_b\Phi(b).$$

One should realize that CN_p is independent of the actual choice for Φ and that CN_p does not have to be greater than 1. In order to have a more workable quantity, it was suggested in [15] to use the following estimate for CN_p,

$$(5.21) \qquad\qquad \gamma_p = \|Q^{-1}\|_p.$$

This estimate is meaningful only if we make the following (not restrictive) assumption. Let $\Phi(t) = T(t)D(t)$ be such that
    (i) $\max_{a \le t \le b}\|d_{jj}(t)\|_p = \max_{a \le s \le b}\|d_l(s)\|_p$, $\forall j, l$;
    (ii) $\max_{a \le t \le b}\|\Phi(t)\|_p = 1$.
We obtain
THEOREM 23.

$$\frac{1}{\max_{a \le t \le b}\|S^T(t)\|_p} \frac{1}{\max_{a \le t \le b}\|T(t)\|_p}\left(\frac{1}{n}\right)^{1/p}\gamma_p \le CN_p \le \gamma_p.$$

Proof. The second inequality is an immediate consequence of our normalization assumption and the fact that $\|\Phi(t)Q^{-1}\|_p \le \|\Phi(t)\|_p\|Q^{-1}\|_p$. To show the first inequality, let z be a maximizing vector of Q^{-1}, i.e. $\|Q^{-1}z\|_p = \gamma_p\|z\|_p$. Define $y := Q^{-1}z$; then

$$(5.22) \qquad \|\Phi(t)Q^{-1}z\|_p = \|T(t)D(t)y\|_p \ge glb_p(T(t))\|\phi(t)y\|_p.$$

Now we have

$$(5.23) \qquad \max_t\|D(t)y\|_p \ge \max_t|d_{ii}(t)|\cdot\|y\|_\infty \ge \max_t|d_{ii}(t)|\left(\frac{1}{n}\right)^{1/p}\|y\|_\infty$$

(where i is arbitrary). Finally, from $D(t) = T^{-1}(t)\Phi(t)$, we derive

$$(5.24) \qquad \max_t|d_{ii}(t)| \ge \max_t\left\{glb_p(T^{-1}(t))\|\Phi(t)\|_p\right\}.$$

Substituting (5.24) into (5.23) and this into (5.22) where we now take the max over all t yields

$$CN_p \ge \left[\min_t glb_p(T^{-1}(t))\right]\left[\min_t glb_p(T(t))\right]\left(\frac{1}{n}\right)^{1/p}\gamma_p\frac{\|z\|_p}{\|z\|_p},$$

since $glb_p(T^{-1}) = 1/\|T\|_p$ and $glb_p(T) = 1/\|T^{-1}\|_p$ the result immediately follows. □

Note that whereas CN_p is independent of the choice of Φ, γ_p is not. It appears that γ_p is a sharper estimate the less "skew" the direction matrix T is. For a useful estimate γ_p, we therefore have to choose a Φ such that the basis solutions have fairly well separated directions (if this is at all possible). The preceding turning point problem provides a nice example to demonstrate this. To this end, let $[a,b] = [-1,1]$ and assume

that ε is sufficiently small to let the asymptotic behavior ($t \to \pm\infty$) be valid already for $t \to \pm 1$. As boundary condition we consider

$$(5.25) \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x(-1) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x(1) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(here $x = (u, \sqrt{\pi\varepsilon}\, u')^T$, cf. (5.12)).

Choosing Φ according to (5.14) results in

$$(5.26) \qquad Q \approx \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Hence $\gamma_\infty \approx 2$. Although this bound seems small, it may not be a sharp estimate for CN_∞, for as we can see from (5.14) we have $\max_t \|S(t)\|_\infty \sim e^{1/\varepsilon}$, implying that the lower bound in Theorem 23 tends to zero as $\varepsilon \to 0$. On the other hand, if we choose ΦM as in (5.14'), we obtain

$$(5.26) \qquad Q \approx \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence $\gamma_\infty \approx 1$. Moreover, it follows from (5.15) (see also Table 1) that $\max_t \|S^T(t)\|_\infty \cdot \max_t \|T(t)\|_\infty \approx 3$. Hence

$$(5.27) \qquad \frac{1}{3} \lesssim CN_\infty \lesssim 1.$$

REFERENCES

[1] F. BAUER, J. STOER, AND C. WITZGALL, *Absolute and monotonic norms*, Numer. Math., 3 (1961), pp. 257–264.

[2] R. BELLMAN, *Stability Theory of Differential Equations*, McGraw-Hill, New York, 1965.

[3] R. BULIRSCH AND J. STOER, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1980.

[4] L. CESARI, *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations*, Academic Press, New York, 1963.

[5] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[6] W. A. COPPEL, *Stability and Asymptotic Behavior of Differential Equations*, Heath, Boston, 1965.

[7] ———, *Dichotomies in Stability Theory*, Springer-Verlag, New York, 1978.

[8] G. DAHLQUIST, *Stability and error bounds in the numerical integration of ordinary differential equations*, Trans. Royal Inst. of Tech., Stockholm, 1959.

[9] G. DAHLQUIST, L. EDSBERG, G. SKÖLLERMO AND G. SÖDERLIND, *Are the numerical methods and software satisfactory for chemical kinematics*? in Numerical Integration of Differential Equations and Large Linear Systems, J. Hinze, ed., Proceedings, Bielefeld 1980, Springer-Verlag, New York, 1982.

[10] F. R. GANTMACHER, *The Theory of Matrices*, Vol. 2, Chelsea, New York, 1959.

[11] E. J. HARRIS AND J. F. MILES, *Stability of Linear Systems: Some Aspects of Kinematic Similarity*, Academic Press, New York, 1980.

[12] B. KÅGSTRÖM, *Bounds and perturbations bounds for the matrix exponential*, BIT, 17 (1977), pp. 39–57.

[13] H. B. KELLER AND J. B. KELLER, *Exponential-like solution of systems of linear ordinary differential equations*, J. Soc. Ind. Appl. Math, 10 (1962), pp. 246–259.

[14] B. LINDBERG, *A dangerous property of methods for stiff differential equations*, BIT, 14 (1974), pp. 430–436.

[15] R. M. M. MATTHEIJ, *The conditioning of linear boundary value problems*, SIAM J. Numer. Anal., 19 (1982), pp. 963–978.

[16] ———, *Estimates for the errors in the solutions of linear boundary value problems, due to perturbations*, Computing, 27 (1981), pp. 299–318.

[17] V. V. NEMYTSKIJ AND V. V. STEPANOV, *Qualitative Theory of Differential Equations*, Princeton Univ. Press, Princeton, NJ, 1960.

[18] A. VAN DER SLUIS, *Estimating the solutions of slowly varying differential equations*, Preprint 152, Dept. Mathematics, Univ. Utrecht, 1980.

[19] G. SÖDERLIND, *On nonlinear difference and differential equations*, BIT, 24 (1984), to appear.

[20] T. STRÖM, *On logarithmic norms*, SIAM J. Numer. Anal., 2 (1975), pp. 741–753.

[21] W. WASOW, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley Interscience, New York, 1965.

[22] J. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.