Journal of Computational Mathematics, Vol.17, No.5, 1999, 457–462.

RELATIONS BETWEEN TWO SETS OF FUNCTIONS DEFINED BY THE TWO INTERRELATED ONE-SIDE LIPSCHITZ CONDITIONS*1)

Shuang-suo Zhao
(*Department of Mathematics, Lanzhou University, Lanzhou 730000, China*)

Chang-yin Wang
(*Communication center, Department of Communications, Gansu Province, Lanzhou 730030, China*)

Guo-feng Zhang
(*Department of Mathematics, Lanzhou University, Lanzhou 730000, China*)

Abstract

In the theoretical study of numerical solution of stiff ODEs, it usually assumes that the righthand function $f(y)$ satisfy one-side Lipschitz condition

$$< f(y) - f(z), y - z > \leq \nu' \|y - z\|^2, f : \Omega \subseteq C^m \to C^m,$$

or another related one-side Lipschitz condition

$$[F(Y) - F(Z), Y - Z]_D \leq \nu'' \|Y - Z\|_D^2, F : \Omega^s \subseteq C^{ms} \to C^{ms},$$

this paper demonstrates that the difference of the two sets of all functions satisfying the above two conditions respectively is at most that $\nu' - \nu''$ only is constant independent of stiffness of function f.

Key words: Stiff ODEs, One-side Lipschitz condition, Logarithmic norm.

In the theoretical study of numerical solution of stiff ODEs, authors usually assume that the righthand function f of

$$y'(t) = f(y(t)), \quad y(t_0) = y_0, \quad t \in [t_0, T], \quad f : \Omega \subseteq C^m \to C^m, \tag{1}$$

satisfy the one-side Lipschitz condition[1,2,3]

$$< f(y) - f(z), y - z > \leq \nu \|y - z\|^2, \forall y, z \in \Omega, \tag{2}$$

* Received February 27, 1995.
1)Supported by the national natural science foundation.

however, in some cases(such as study of existence and uniqueness of the solution), the function f is assumed to satisfy another one-side Lipschitz condition

$$[F(Y)-F(Z),Y-Z]_D \leq \nu\|Y-Z\|_D^2, \tag{3}$$

where Ω is a convex domain in C^m, $Y = (y_1^T, y_2^T, \cdots, y_s^T)^T \in \Omega^s := \overbrace{\Omega\times\Omega\times\cdots\times\Omega}^{s\ \text{times}}$, $F(Y)=(f^T(y_1), f^T(y_2),\cdots,f^T(y_s))^T$, $<\cdot,\cdot>$ is an inner-product in C^m, $\|\cdot\|$ is the corresponding norm, $D=(d_{ij})$is a s-by-s Hermite positive definite matrix, $[F(Y),Z]_D = \sum_{i,j=1}^s d_{ij} < f(y_i), z_j >$, $\|\cdot\|_D$ is the corresponding norm.

Definition:

$$\mathcal{F}_1(\nu) = \{f(y) \mid Re < f(y)-f(z), y-z > \leq \nu\|y-z\|^2,\ f'(y)\ is\ existed, \forall y,z\in\Omega\},$$

$$\mathcal{F}_2(\nu) = \{f(y) \mid Re[F(Y)-F(Z),Y-Z]_D \leq \nu\|Y-Z\|_D^2, f'(y)\ is\ existed, \forall Y,Z\in\Omega^s\},$$

where $f'(y)$ is a Frechet-derivative of $f(y)$ with respect to y. Up to date, there is no result for the relation of $\mathcal{F}_1(\nu)$ and $\mathcal{F}_2(\nu)$. The goal of this paper is to investigate this problem.

Theorem 1. *If D is a diagonally positive definite matrix, then*

$$\mathcal{F}_1(\nu) = \mathcal{F}_2(\nu).$$

Proof. For $\forall f(y)\in\mathcal{F}_2(\nu)$, it follows from the definition that

$$Re\sum_{i=1}^s d_{ii} < f(y_i)-f(z_i), y_i-z_i >= Re[F(Y)-F(Z),Y-Z]_D \leq \nu\|Y-Z\|_D^2, \tag{4}$$

if $f(y)\notin\mathcal{F}_1(\nu)$, then there exist $y,z\in\Omega$ such that

$$Re < f(y)-f(z), y-z >> \nu\|y-z\|^2.$$

Let $Y=(y^T,y^T,\cdots,y^T)^T$ and $Z=(z^T,z^T,\cdots,z^T)^T\in\Omega^s$, then

$$Re\sum_{i=1}^s d_{ii} < f(y)-f(z), y-z >> \nu\|Y-Z\|_D^2.$$

That is conflict with (4), so $\mathcal{F}_2(\nu)\subseteq\mathcal{F}_1(\nu)$. On the other hand, it is obvious that $\mathcal{F}_1(\nu)\subseteq\mathcal{F}_2(\nu)$. Therefore, $\mathcal{F}_1(\nu)=\mathcal{F}_2(\nu)$.

Theorem 2. *Assume that the D be a Hermite positive definite matrix and $f(y) = By+\hat{B}$ be a linear function, then $f\in\mathcal{F}_1(\nu) \iff f\in\mathcal{F}_2(\nu)$.*

Proof. For the inner-products $<\cdot,\cdot>$ and standard inner-product $(y,z)=y^*z$ in C^m, there exists a Hermite positive definite matrix Q such that

$$< y,z >= (y,Qz), \quad \forall y,z\in C^m.$$

So, for an arbitrary block diagonal matrix H=block-diag$(B, B, \cdots, B) \in C^{ms\times ms}$, we have

$$[HY, Z]_D = (HY, (D \otimes Q)Z) = (GHY, GZ),\ \forall Y, Z \in C^{ms},$$

where $G = (D \otimes Q)^{\frac{1}{2}}$, $\otimes$ is Kronecker product symbol. Especially, when $Z = Y$, we have

$$[HY, Y]_D = (GHY, GY), \quad [Y, Y]_D = (GY, GY),\ \forall Y \in C^{ms}.$$

It is easy to conclude that

$$Re\frac{[HY, Y]_D}{[Y, Y]_D} = Re\frac{(GHG^{-1}Z, Z)}{(Z, Z)} = \frac{1}{2}\frac{((GHG^{-1} + G^{-1}H^*G)Z, Z)}{(Z, Z)}, Z = GY, \qquad (6)$$

when $f(y) = By + \hat{B}$, $F(Y) - F(Z) = H(Y - Z)$, where $H = I_s \otimes B$, I_s is a s-by-s identity matrix. It is obvious that

$$\begin{cases} GHG^{-1} = (D \otimes Q)^{\frac{1}{2}}(I_s \otimes B)(D \otimes Q)^{-\frac{1}{2}} = I_s \otimes (Q^{\frac{1}{2}}BQ^{-\frac{1}{2}}), \\ G^{-1}H^*G = I_s \otimes (Q^{-\frac{1}{2}}B^*Q^{\frac{1}{2}}). \end{cases} \qquad (7)$$

It folows from(6) and (7) that

$$\begin{aligned} Re\frac{[F(Y) - F(Z), Y - Z]_D}{[Y - Z, Y - Z]_D} &= Re\frac{[H(Y - Z), Y - Z]_D}{[Y - Z, Y - Z]_D} \\ &= \frac{1}{2}\frac{((I_s \otimes (Q^{\frac{1}{2}}BQ^{-\frac{1}{2}} + Q^{-\frac{1}{2}}B^*Q^{\frac{1}{2}}))\tilde{Z}, \tilde{Z})}{(\tilde{Z}, \tilde{Z})}, \quad \tilde{Z} = G(Y - Z). \end{aligned}$$

For $Q^{\frac{1}{2}}BQ^{-\frac{1}{2}} + Q^{-\frac{1}{2}}B^*Q^{\frac{1}{2}}$ is a Hermite matrix, so,

$$\max_{Y\neq Z} Re\frac{[F(Y) - F(Z), Y - Z]_D}{[Y - Z, Y - Z]_D} = \frac{1}{2}\lambda_{\max}(Q^{\frac{1}{2}}BQ^{-\frac{1}{2}} + Q^{-\frac{1}{2}}B^*Q^{\frac{1}{2}}). \qquad (8)$$

On the other hand, we have also

$$\begin{aligned} \max_{y\neq z}\{Re\frac{< f(y) - f(z), y - z >}{< y - z, y - z >}\} &= \max_{y\neq z}\{Re\frac{< B(y - z), y - z >}{< y - z, y - z >}\} \\ &= \frac{1}{2}\lambda_{\max}(Q^{\frac{1}{2}}BQ^{-\frac{1}{2}} + Q^{-\frac{1}{2}}B^*Q^{\frac{1}{2}}), \end{aligned}$$

compared with (8), the desired result holds.

Lemma. *If $f(y) \in \mathcal{F}_1(\nu)$, then $\mu(f'(z)) \leq \nu, \forall z \in \Omega$; if $f(y) \in \mathcal{F}_2(\nu)$, then $\mu(F'(Y)) \leq \nu, \forall Y \in \Omega^s$, where $\mu(A)$ is the logarithmic norm of n-by-n complex matrix A, namely,*

$$\mu(A) = \max_{z\in C^n, z\neq 0} Re\frac{[Az, z]}{[z, z]}, \quad n = m, or, n = ms,$$

$[\cdot, \cdot]$ is the inner-product in C^n.

Proof. If $f(y) \in \mathcal{F}_1(\nu)$, then

$$Re < f(y) - f(z), y - z > \leq \nu \|y - z\|^2, \forall y, z \in \Omega.$$

Let $y = z + tw, w \in C^m, t \in R, z \in \Omega$, for the Ω is a convex domain, so $y \in \Omega$ as t is small enough, from the above inequality, we have

$$Re < f(z + tw) - f(z), tw > \leq \nu t^2 \|w\|^2.$$

It follows that

$$Re < f'(z)w, w > \leq \nu \|w\|^2, \forall z \in \Omega, \forall w \in C^m.$$

This showes that $\mu(f'(z)) \leq \nu$. The proof of the another part is similar.

Theorem 3. *Assume that the D be a Hermite positive definite matrix, $f(y)$ satisfy*

$$\|f'(y) - f'(z)\| \leq M\|y - z\|, \qquad \forall y, z \in \Omega, \tag{9}$$

then i) $f(y) \in \mathcal{F}_2(\nu + \nu')$ as $f(y) \in \mathcal{F}_1(\nu)$,

ii) $f(y) \in \mathcal{F}_1(\nu + \nu'')$ as $f(y) \in \mathcal{F}_2(\nu)$,

where ν', ν'' are defined in (11), they are only dependent on the D, $< \cdot, \cdot >$, M and Ω, and independent of stiffness of function f.

Proof. For $\forall Y_i = (y_{i1}^T, y_{i2}^T, \cdots, y_{is}^T)^T \in \Omega^s$(i=1,2), we have

$$F(Y_1) - F(Y_2) = H(Y_1 - Y_2),$$

where H=block-diag$(B_1, B_2, \cdots, B_s)$, $B_j = \int_0^1 f'(y_{2j} + \theta(y_{1j} - y_{2j}))\, d\theta$, $j = 1(1)s$. Let $H_0 = I_s \otimes B_1$, H_1=block-diag$(0, B_2\text{-}B_1, \cdots, B_s\text{-}B_1)$, then

$$H \equiv H_0 + H_1, \ \forall Y_1, Y_2 \in \Omega^s.$$

Therefore,

$$[HW, W]_D = [H_0W, W]_D + [H_1W, W]_D, \ \forall Y_1, Y_2 \in \Omega^s, \forall W \in C^{ms}. \tag{10}$$

Definition:

$$\nu' = \max_{Y_1, Y_2 \in \Omega^s} \max_{W \neq 0} Re(\frac{[H_1W, W]_D}{[W, W]_D}), \ \nu'' = \max_{Y_1, Y_2 \in \Omega^s} \max_{W \neq 0} Re(\frac{[-H_1W, W]_D}{[W, W]_D}). \tag{11}$$

It is obvious for $\forall Y_1, Y_2 \in \Omega^s, \forall W \in C^{ms}$ that

$$Re[H_1W, W]_D \leq \nu'[W, W]_D, \ Re[-H_1W, W]_D \leq \nu''[W, W]_D. \tag{12}$$

For the arbitrarily fixed $Y_1, Y_2 \in \Omega^s$, following the proving of the theorem 2, we have

$$\max_{W \neq 0} Re \frac{[H_0W, W]_D}{[W, W]_D} = \max_{w \neq 0} Re \frac{< B_1w, w >}{< w, w >}. \tag{13}$$

It is obvious that

$$\max_{w\neq 0} Re\frac{< B_1 w, w >}{< w, w >} = \max_{w\neq 0}\int_0^1 Re\frac{< f'(y_{21}+\theta(y_{21}-y_{22}))w, w >}{< w, w >}\, d\theta.$$

If $f\in\mathcal{F}_1(\nu)$, from the lemma, we have

$$\max_{w\neq 0} Re\frac{< B_1 w, w >}{< w, w >}\leq \nu.$$

By the above inequality and (13), we have

$$Re[H_0W,W]_D\leq \nu\|W\|_D^2,\ \forall Y_1,Y_2\in\Omega^s,\forall W\in\ C^{ms}. \tag{14}$$

Let $W=Y_1-Y_2$, it follows from (10),(12) and (14) that

$$Re[H(Y_1-Y_2),Y_1-Y_2]_D\leq(\nu+\nu')\|Y_1-Y_2\|_D^2,\ \forall Y_1,Y_2\in\Omega^s,$$

this indicates $f(y)\in\mathcal{F}_2(\nu+\nu')$.

If $f\in\ \mathcal{F}_2(\nu)$, from (10),(12) and the lemma, we have

$$Re[H_0W,W]_D=Re[HW,W]_D+Re[-H_1W,W]_D\leq(\nu+\nu'')\|W\|_D^2.$$

Using (13), we obtain

$$< B_1w,w >\leq(\nu+\nu'')\|w\|^2,\ \forall y_{11},y_{21}\in\Omega,\forall w\in C^m. \tag{15}$$

Let $w=y_{11}-y_{21},y_{11}=y,y_{21}=z$, we obtain from (14)

$$< f(y)-f(z),y-z >\leq(\nu+\nu'')\|y-z\|^2,\forall y,z\in\Omega.$$

This showes $f(y)\in\mathcal{F}_1(\nu+\nu'')$.

Finally, we evaluate ν' and ν'', from (11), (9) and the definition of H_1, it follows that

$$\begin{aligned}\max(|\nu'|,|\nu''|)&\leq\max_{Y_1,Y_2\in\Omega^s}\|H_1\|_D=\max_{Y_1,Y_2\in\Omega^s}\max_{Y\neq 0}\left(\frac{[H_1Y,H_1Y]_D}{[Y,Y]_D}\right)^{\frac{1}{2}}\\&=\max_{Y_1,Y_2\in\Omega^s}\max_{Z\neq 0}\frac{(GH_1G^{-1}Z,GH_1G^{-1}Z)^{\frac{1}{2}}}{(Z,Z)^{\frac{1}{2}}}\\&\leq|G|_{ms}|G^{-1}|_{ms}\max_{Y_1,Y_2\in\Omega^s}|H_1|_{ms}\\&=|G|_{ms}|G^{-1}|_{ms}\|Q^{\frac{1}{2}}\|\cdot\|Q^{-\frac{1}{2}}\|\max_{2\leq j\leq s}\max_{y_{1j},y_{2j}\in\Omega}\|B_j-B_1\|\\&\leq 3|G|_{ms}|G^{-1}|_{ms}\|Q^{\frac{1}{2}}\|\cdot\|Q^{-\frac{1}{2}}\|M\rho(\Omega),\end{aligned}$$

where $|\cdot|_{ms}$ denotes the spectral norm in C^{ms}, $\rho(\Omega)$ is the diameter of the set Ω. Obviously $\max(|\nu'|,|\nu''|)\to 0$ as $\rho(\Omega)\to 0$. It follows that when the D is a nondiagonal positive definite matrix, if $\rho(\Omega)$ is very small, then the difference of $\mathcal{F}_1(\nu)$ and $\mathcal{F}_2(\nu)$ is also very small.

References

[1] K. Dekker, J.G. Verwer , Stability of Runge-Kutta Methods for Stiff Nonlinear Differential equations, Amsterdam · New York · Oxford, North-Holland, 1984.

[2] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations, Runge-Kutta and General Linear Methods, Chichester · New York · Brisbane · Toronto · Singapore, John Wiley & Sons, 1987.

[3] Jiao-xun Kuang, Jia-xiang Xiang, On the D-suitability of implicit Runge-Kutta methods, *BIT,* **29** (1989), 321-327.

[4] Shou-fu Li, Existence and Uniqueness of solution of a class of operator equations, *Sci. Bulle.(in chinese),* **37**:5 (1992), 388-391.