# FU JEN STUDIES

## SCIENCE AND ENGINEERING

NO. 31, DEC. 1997

## CONTENTS

# 輔仁學誌—理工類

中華民國八十六年十二月　　第三十一期

## 目　錄

# On The Stabilizing Uncertain Lur'e-Postnikov Systems by Variable Structure Control

KOU-CHENG HSU
*Department of Electronic Engineering*
*Fu-Jen Catholic University*
*Hsin-Chuang, Taipei 24205*

## ABSTRACT

This paper proposes a robust control method for the uncertain Lur'e-Postnikov systems. It shows that, in the sliding mode, the uncertain system with "nonlinear" input possesses the insensitivity property to the uncertainties and disturbances just as those with "linear" inputs do. Then a robust variable structure control law is investigated such that the trajectories of uncertain Lur'e-Postnikov system can be forced onto the sliding mode. The required informations about uncertain dynamics in the system are that the uncertainties are bounded by known functions and the property of the "nonlinear" inputs. Furthermore, the sliding mode can be designed to converge within a specified exponential speed.

***Key Words***: Lur'e-Postnikov system, variable structure control, robust control, nonlinear input, series nonlinearity.

## INTRODUCTION

The subject of robust control deals with the design of control systems subjected to uncertainties and/or external disturbances. From the standpoint of design, the uncertain system with linear inputs is more preferable than those with nonlinear inputs because of its simplicity. Most of the robust control are devoted to the stabilization problem of this kind of system[1-8]. The assumption of linear inputs is that the system

model is indeed linearizable. However, in control systems, there are many nonlinearities in the control inputs and their effects can not be derived from linear methods. Those nonlinearities inherently arise from practical actutors in system realization, for example, saturation, quantization, backlash, deadzone, and so on. The existence of input nonlinearities is a source of degradation or, even worse, instability of system performance. Consequently, the problem of stability analysis of control system design accounting for input nonlinearities has been a newly concerned research area[9-13] (also see the numerous references cited therein). In this case, the existing works[1-8] can not be applied to stabilize the uncertain systems with nonlinear inputs. In addition to the input nonlinearities, we have to face with plant uncertainties originated from various sources, such as variation of plant parameters, inaccuracy from identification, *etc*. Therefore, in spite of the existed works discussing about the robust control, it is necessary to develop an effective robust control method for the uncertain systems with nonlinear inputs, because the problem of stability analysis of robust control systems with plant uncertainties is as important as that with input nonlinearities.

Among the nonlinear input systems, Lur'e-Postnikov system is a well known typical example. Many physical systems can be naturally interpreted as consisting of a "linear part" and a "nonlinear part," so that Lur'e-Postnikov system is, reasonably, taken as a general system model[14,15]. Therefore, in this paper, an effective robust control method is developed to guarantee the asymptotic stability of uncertain Lur'e-Postnikov system.

For the robust control, variable structure control (VSC) is quite popularly utilized since the variable structure control systems possess several attractive advantages in sliding mode, for example, fast response, excellent transient performance, and insensitivity to the variation of plant parameters or external disturbances[16,17]. However, most of the existing works about the VSC systems concentrated on the systems with "linear" inputs[16-24].

To cope with the stabilization problem of uncertain Lur'e-Postnikov systems and to make the controlled uncertain Lur'e-Postnikov systems be insensitive to uncertain parameters and external disturbances, in this paper, a VSC method is developed for

the purpose of stabilizing the uncertain Lur'e-Postnikov systems. To achieve such goal, some properties of the Lur'e-Postnikov systems are firstly reviewed. Then a robust variable structure controller is derived based on these properties. Furthermore, the developed variable structure controller is modified to ensure the convergence speed of sliding mode within a specified exponential speed. Finally, a numerical example is illustrated to examine the validity of the proposed VSC controller.

## SYSTEM DESCRIPTION AND PROBLEM FORMULATION

A general description of an uncertain Lur'e-Postnikov system[9] is expressed in the form as

$$\dot{x}(t) = Ax(t) + B\Phi(u) + Be(x,u,p,t) \tag{1}$$

where $x(t) \in R^n$ is the state variable, $u(t) \in R^m$ is the control input, $p(t) \in R^q \subset \Sigma$ is the uncertain parameter, and $e(t) \in R^s$ is the external disturbance. $A \in R^{n\times n}$ is the state matrix, $B \in R^{n\times m}$ is the input matrix, and $\Phi: R^m \to R^m$ is a continuous function such that $\Phi(0) = 0$. It is also assumed that for any initial condition $x(t_0) = x_0 \in R^n$ at $t = t_0$, parameter $p \in R^q$, and control input $u(t)$, there exists a unique solution $x(t, x_0, p, u)$ satisfying equation (1).

Through this paper, two standard assumptions regarding system (1) are made:

**Assumption 1**: For the nominal part of the uncertain Lur'e-Postnikov system described in eq. (1), matrix $A$ and matrix $B$ is a controllable pair.

**Assumption 2**: For the lumped uncertain terms of the system, or $e(x,u,p,t)$, there exist known non-negative constants $k_1$, $k_2$, and $k_3$, such that

$$\| e(x,u,p,t) \| \leq k_1 \| x \| + k_2 \| u \| + k_3$$
$$\forall (x,u,p,t) \in R^n \times R^m \times R^q \times R \tag{2}$$

In equation (1), the nonlinear inputs applied to the system satisfy the following property:

$$ru^T u \leq u^T \Phi(u) \tag{3}$$

where $r$ is a positive nonzero constant, $u^T$ is the transpose of $u$, and $\Phi(u)$ satisfies $\Phi(0) = 0$.

For the uncertain nonlinear system, the main theme of this paper is to derive an effective VSC method such that the controlled uncertain Lur'e-Postnikov system is stable and insensitive to the system uncertainties. In the following sections of this article, the notation $\|\cdot\|$ denotes the usual Euclidean norm or the corresponding induced matrix norm.

## VARIABLE STRUCTURE CONTROL DESIGN

Usually the VSC design is a two-stage process. The first phase is to choose a set of switching surfaces so that the original system, restricted to the intersection of the switching surfaces (sliding modes), results in the desired behavior. The second phase is to determine a switching control that is able to force the trajectory of the system approaching to and staying on the sliding surface[16-24]. Hereby, the switching surfaces are firstly defined as

$$S(t) = Cx(t) = 0 \tag{4}$$

where $C \in R^{m\times n}$ is a constant matrix which requires a nonzero determinant of the product matrix $(CB)$, that is, $det(CB) \neq 0$. Because most of the existed studies related to VSC[18-24] do not discuss the system with nonlinear input, it is necessary to examine if the property in the sliding mode, described in eq. (4), is valid for the system with nonlinear input, shown in eq. (1). Once the sliding mode $S(t) = 0$ is obtained, it is always accompanied with the condition of $\dot{S}(t) = 0$. Hence, in the sliding mode, the property of the system with nonlinear input can be inspected by inserting eq. (1) into the derivative of eq. (4), or $\dot{S} = 0$, to yield

$$\begin{aligned}\dot{S}(t) &= C\dot{x}(t)\\ &= C[Ax(t) + B\Phi(u) + Be(x,u,p,t)]\\ &= 0\end{aligned} \tag{5}$$

Therefore, the equivalent control $\Phi_{eq}$ in the sliding mode $S = 0$ is given by

$$\Phi_{eq} = -(CB)^{-1}C(Ax + Be) \tag{6}$$

**Remark 1**: It is noted that the equivalent control $\Phi_{eq}$ is only a mathematically derived

tool for the purpose of analyzing a sliding motion rather than a real control law being generated in practical systems. In fact, the equivalent control $\Phi_{eq}$ is only realizable through a nonlinear controller if the system is absent from uncertainties, that is, all uncertainties are zero in the nominal system. It is also noted that the equivalent control generates an "ideal" sliding motion on the switching hyperplane while the real variable structure controller generates a trajectory close to the ideal sliding motion around the switching hyperplane.

Introducing equation (6) into eq. (1) produces the equivalent dynamic system with nonlinear input in the sliding mode as:

$$\begin{aligned} S(t) &= 0 \\ \dot{x}(t) &= Ax(t) + B\Phi_{eq} + Be(x, u, p, t) \\ &= [I - B(CB)^{-1}C]Ax \end{aligned} \tag{7}$$

where $I$ is an $n \times n$ identity matrix. From equation (7), it can be seen that the invariance condition[25] also holds even though the system is with "nonlinear" inputs.

**Remark 2**: From the above analysis, it can be concluded that the uncertain system with "nonlinear" inputs in the sliding mode possesses the same property as those with "linear" inputs in the sliding mode being able to be made. Accordingly, the design of the switching surfaces can be achieved through same deriving procedures for the systems with "linear" inputs.

From the analysis mentioned above, it can be seen that how to drive the system trajectories onto the sliding mode is the key work for system stabilization. Before stating the scheme of the controller, the reaching condition of the sliding mode is given below [23]:

**Lemma 1**: The motion of the sliding mode (6) is asymptotically stable, if the following condition is held

$$S^T(t)\dot{S}(t) < 0, \qquad \forall t \geq 0 \tag{8}$$

To fulfil the condition stated in eq.(8), the desired switching control is suggested by

$$u(t) = -\frac{B^T C^T S}{\| B^T C^T S \|}\phi(x, t) \tag{9}$$

where

$$\phi(x,t) = \frac{\beta}{r-k_2}\left\{[\,\|\,(CB)^{-1}CA\,\|\, + k_1]\,\|\,x\,\|\, + k_3\right\}, \qquad \beta > 1, r > k_2 \quad (10)$$

It notes that $\|\,u(t)\,\| = \phi(x,t)$ is implied by eq. (9). The following theorem shows that the proposed control in eq. (9) drives uncertain nonlinear input system (1) onto the sliding mode $S(t) = 0$.

**Theorem 1**: Consider the uncertain system (1) subjected to Assumptions 1, 2, and inequality of (2). If the input $u(t)$ in equation (1) is given as that indicated by (9), then the system trajectories asymptotically converge to the sliding mode (4).

**Proof**: Consider the reaching condition of the sliding mode (4). If equation (1) is substituted into the derivative of the state $x$ in (8), one can obtain

$$\begin{aligned} S^T(t)\,\dot{S}\,(t) &= S^T C\dot{x}\,(t) \\ &= S^T C[Ax + B\Phi(u) + Be] \end{aligned}$$

Then eq. (2) is applied to the above equation to yield the following inequality expression:

$$\begin{aligned} S^T(t)\,\dot{S}\,(t) \leq &\,\|\,S^TCB\,\|\,\|\,(CB)^{-1}CA\,\|\,\|\,x\,\|\, + S^TCB\Phi(u) \\ &+ \|\,S^TCB\,\|\,(k_1\|\,x\,\|\, + k_2\|\,u\,\|\, + k_3) \end{aligned} \quad (11)$$

From equations (9) and (3), we have

$$\begin{aligned} u^T\Phi(u) &= -\frac{S^TCB}{\|\,B^TC^TS\,\|}\phi(x,t)\Phi(u) \\ &\geq ru^Tu = r\phi^2(x,t) \end{aligned}$$

Therefore, the above expression can be rearranged as

$$\begin{aligned} S^TCB\cdot\Phi(u) &\leq -r\phi^2(x,t)\frac{\|\,B^TC^TS\,\|}{\phi(x,t)} \\ &= -r\phi(x,t)\,\|\,B^TC^TS\,\| \end{aligned} \quad (12)$$

Inserting equation (12) into the right hand side of the inequality in (11), it yields

$$\begin{aligned} S^T(t)\,\dot{S}\,(t) \leq &\,\|\,S^TCB\,\|\,\{\|\,(CB)^{-1}CA\,\|\,\|\,x\,\|\, - r\phi(x,t) \\ &+ (k_1\|\,x\,\|\, + k_2\phi(x,t) + k_3)\} \end{aligned} \quad (13)$$

From the $\phi(x,t)$ defined in eq. (10), the following condition can be derived.

$$S^T(t)\,\dot{S}(t) \leq (1-\beta)\,\|S^TCB\|\,\{[\,\|(CB)^{-1}CA\,\| + k_1]\,\|x\,\| + k_3\} \tag{14}$$

It is conspicuous to result in

$$S^T(t)\,\dot{S}(t) < 0 \tag{15}$$

when $S(t) \neq 0$. Then the proof is completed. □

**Remark 3:** For the case in which $\Phi(u)$ represents a sector-bounded vector, *Theorem* 1 stills provides a sufficient condition to guarantee the system dynamics globally asymptotically stable. Let $\Phi(u(t)) = [\Phi_1(u), \Phi_2(u), \cdots, \Phi_m(u)]^T$ where $\Phi_i(u)$ is characterized by

$$r_1 u_i^2 \leq u_i \Phi_i(u) \leq r_2 u_i^2, \quad \textit{for } i = 1, 2, \cdots, m \tag{16}$$

It can be derived straightforwardly to yield

$$r_1 u^T u \leq u^T \Phi(u) \leq r_2 u^T u \tag{17}$$

where $r_1 = min\{r_{11}, r_{12}, \cdots, r_{1m}\}$ and $r_2 = max\{r_{21}, r_{22}, \cdots, r_{2m}\}$. If we choose $r_1 = r$ and $r_2 \to \infty$, then the last expression satisfies the condition defined in (3), that is, the upper bound of the sector nonlinearities can be released. Therefore, the developed variable structure controller in eq. (9) also works effectively for the systems with sector-bounded input nonlinearities.

**Remark 4:** If $\Phi(u) = u$, which means that the system has no input nonlinearity, then the proposed variable structure control law in (9) is applicable to control such a system as long as $r \leq 1$ is taken since $u^T\Phi(u) = u^Tu \geq ru^Tu$ is always kept.

An effective control method, variable structure control method, has been derived to stabilize the uncertain Lur'e-Postnikov system. It has been shown that the proposed control can drive the uncertain system trajectories onto the sliding mode even though the systems are with "nonlinear" input. Moreover, the uncertain systems can also preserve the invariance condition even though the uncertain systems are subjected to the "nonlinear" inputs.

## CONVERGING RATE CONTROL DESIGN

In the previous section, we have guaranteed that the motion of the sliding mode is

stable under the proposed variable structure control law, but we have not discussed the convergence rate of this motion. Now we discuss this problem in the following theorem.

**Theorem 2:** If the variable structure control of (9) is modified as

$$u = -\frac{B^T C^T S}{\| B^T C^T S \|} \phi_a(x,t) \tag{18}$$

where

$$\phi_a(x,t) = \phi(x,t) + \frac{q}{2(r-k_2)} \frac{S^T S}{\| S^T CB \|} \quad q > 0 \tag{19}$$

where $\phi(x,t)$ and $S(t)$ are the same as defined in eqs. (9) and (4), respectively, and $q$ is a nonzero positive constant. Then the rate of attractiveness to the sliding mode is, at least, as fast as $e^{-qt}$.

**Proof:** Similar to the proof of *Theorem* 1, let theLyapunov function candidate be

$$V(t) = \frac{1}{2} \| S(t) \|^2 \tag{20}$$

Then the following inequality can be derived

$$\begin{aligned} \dot{V}(t) &= S(t)^T \dot{S}(t) \\ &\leq \| S^T CB \| \{ \| (CB)^{-1} CA \| \| x \| - r\phi_a + (k_1 \| x \| + k_2 \phi_a + k_3) \} \\ &\leq (1-\beta) \| S^T CB \| \{ [ \| (CB)^{-1} CA \| + k_1 ] \| x \| + k_3 \} - q \frac{1}{2} S^T S \\ &\leq -q \frac{1}{2} \| S \|^2 = -qV \end{aligned} \tag{21}$$

Multiply $e^{qt}$ on the both sides of (21) and rearrange it to yield

$$\frac{d}{dt}(Ve^{qt}) \leq 0 \tag{22}$$

Integrating eq. (22) between the time interval $[t_0, t]$, we obtain

$$V(t)e^{qt} - V(t_0)e^{qt_0} \leq 0 \tag{23}$$

or

$$V(t) \leq V(t_0)e^{-q(t-t_0)} \tag{24}$$

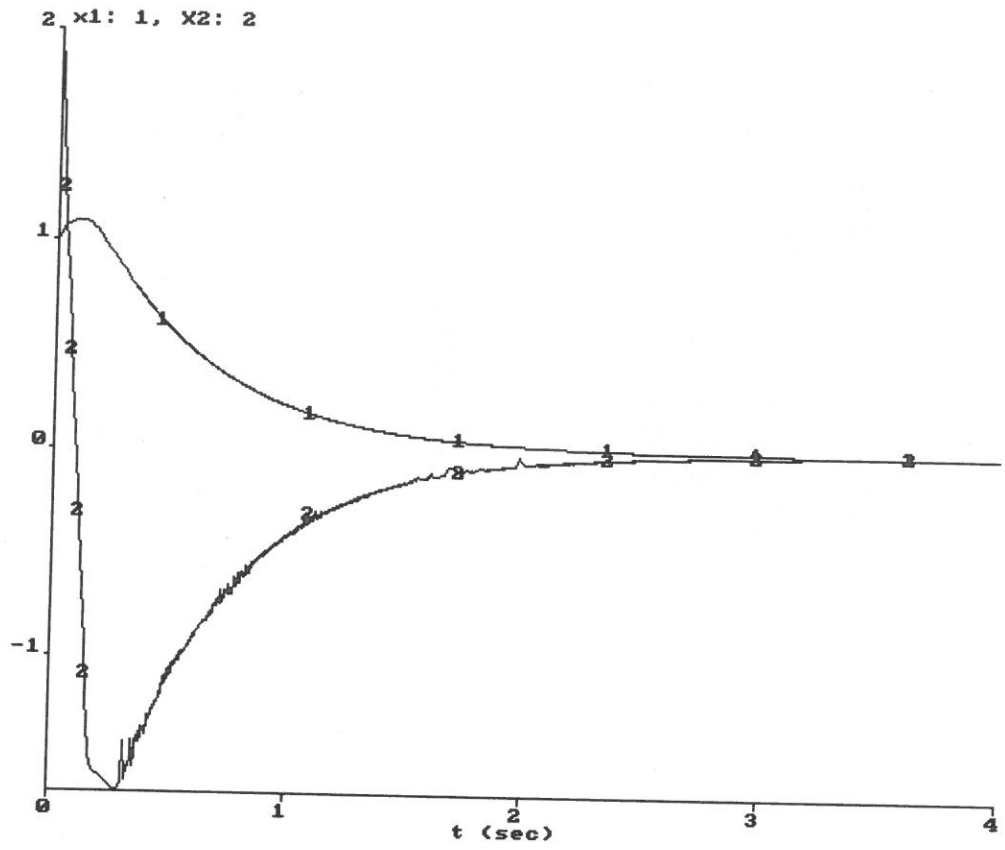



**Fig. 1. State variable dynamics for the system under VSC: $x_1$ and $x_2$**

Because $V(t)$ is the sum of the quadratic form of sliding mode motions, the proof of this theorem is completed. □

In the following section, a numerical example is illustrated to demonstrate the validity of the proposed variable structure control method.

## AN ILLUSTRATIVE EXAMPLE

Consider a Lur'e-Postnikov system which is described by equation (1). It is reshown below

$$\dot{x}(t) = Ax(t) + B\Phi(u) + Be(x, u, p, t)$$

The corresponding parameters of the illustrated system are given as follows:

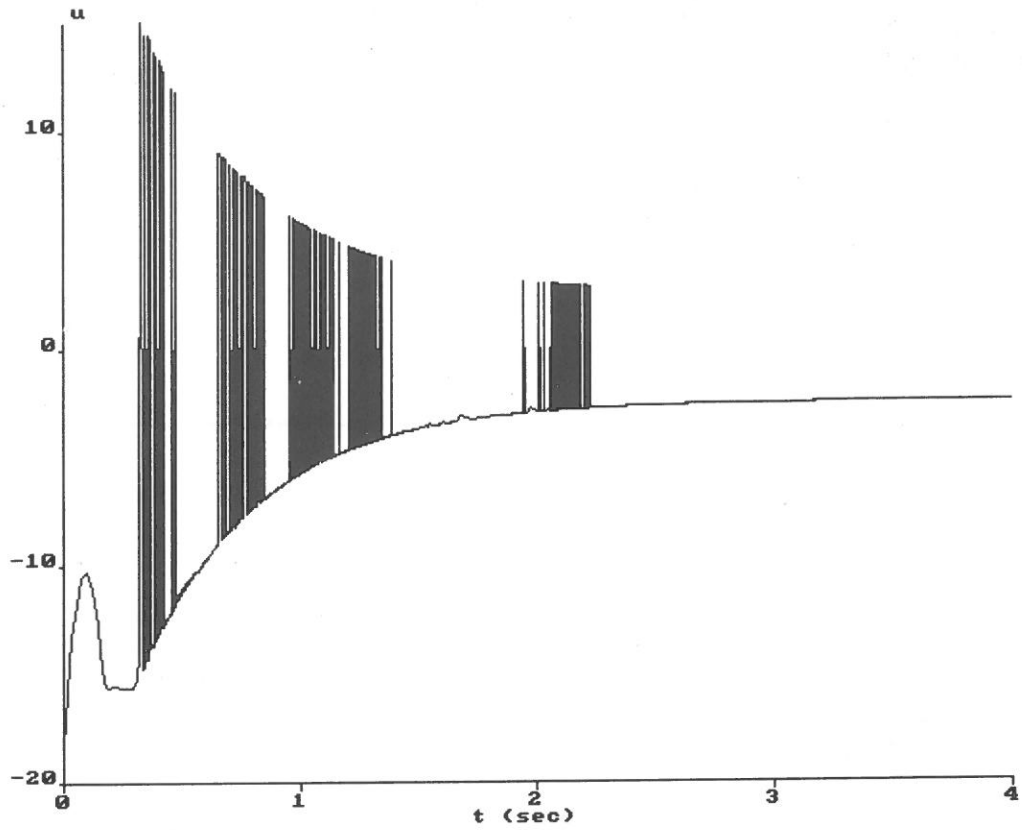



**Fig. 2. The proposed control signal $u(t)$.**

$$A = \begin{bmatrix} 0.0 & 1.0 \\ -0.5 & -1.0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\Phi(u) = (de^{|\sin u|} + \gamma \cos u)u \qquad d > \gamma > 0$$

$$e(x, u, p, t) = l_1(x_1^2 + x_2^2)^{\frac{1}{2}} e^{(1+\sin x_1)} + l_2 \| u \| + l_3 \cos x_2$$

From equation (2), we have the following relation:

$$\begin{aligned} \| e \| &\leq \| l_1 e^{(1+\sin x_1)} \| \, \| x \| + \| l_2 \| \, \| u \| + \| l_3 \cos x_2 \| \\ &\leq k_1 \| x \| + k_2 \| u \| + k_3 \end{aligned}$$

Therefore, we can take $k_i$ as

$$k_1 = \| l_1 \| e^2 \geq \| l_1 e^{(1+\sin x_1)} \|$$

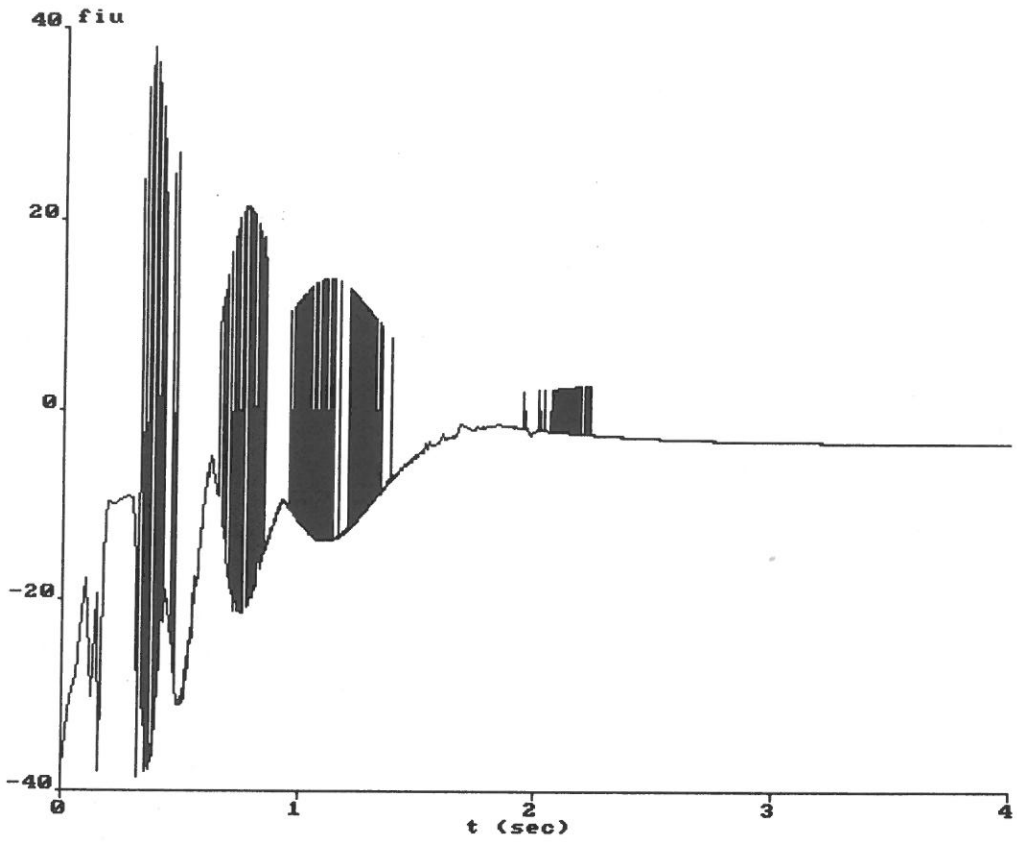



**Fig. 3. The corresponding input function $\Phi(u)$.**

$$k_2 = \| l_2 \| \geq \| l_2 \|$$
$$k_3 = \| l_3 \| \geq \| l_3 \cos x_2 \|$$

Equation (3) yields

$$u^T \Phi(u) = (de^{|\sin u|} + \gamma \cos u) u^2$$
$$\geq (d-\gamma) u^2 = ru^2$$

to result in $r = d - \gamma$. For numerical simulation, we have $d = 1$, $\gamma = 0.5$, $l_1 = 0.04$, $l_2 = 0.2$, and $l_3 = -0.5$ to yield

$$r = 0.5, \quad k_1 = 0.3, \quad k_2 = 0.2, \quad k_3 = 0.5$$

In addition, $\beta = 1.5$ is taken for eq.(14), and $q = 5$ is chosen for eq. (19). $C =$

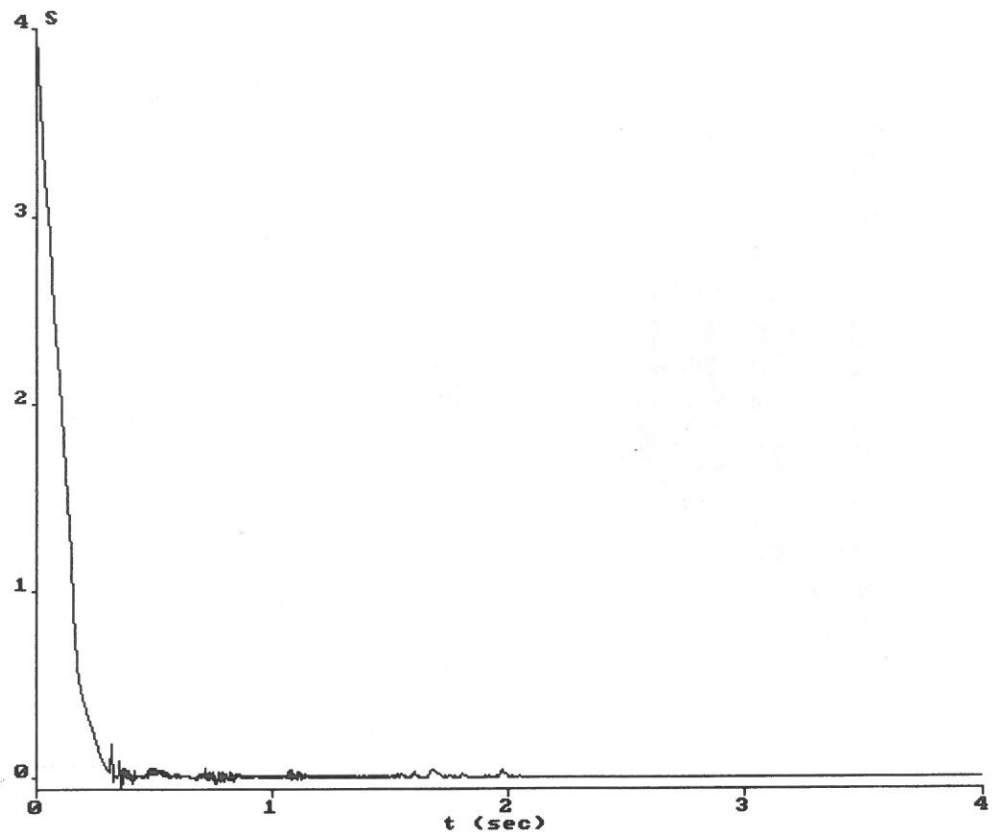



**Fig. 4. The corresponding sliding mode $S(t)$.**

[2 1] is selected for the switching surface and the initial value of $x$ is assumed $[x_1 \quad x_2]^T = [1 \quad 2]^T$ in simulation. The computer simulation for this system has been performed with sampling step of 0.01 *sec*. Fig. 1 shows the transient response of $x_1$ and $x_2$. Fig. 2 demonstrates the proposed control $u(t)$ which forces the system to the designed sliding mode $S(t)$. The corresponding nonlinear input function $\Phi(u)$ is illustrated in Fig. 3. Fig. 4 gives the corresponding sliding mode $S(t)$. Fig. 5 gives the phase plane between $x_1$ and $x_2$. The corresponding converging speed control is shown in Fig. 6, where the real converging rate of $V(t)$ is within the proposed speed range.

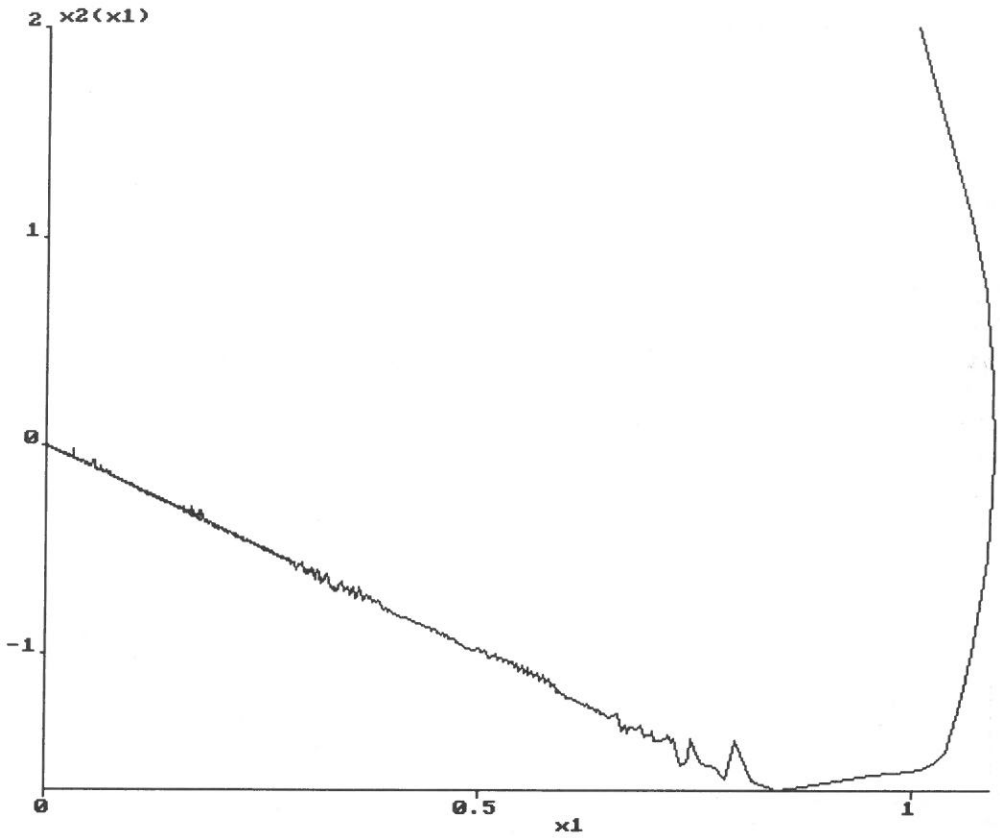



**Fig. 5. The phase plane between $x_1$ *and* $x_2$: $x_2$ $(x_1)$.**

## CONCLUSIONS

An effective control method is proposed for stabilizing uncertain Lur'e-Postnikov systems. The main contribution of the current work is the construction of variable structure controller for uncertain systems with "nonlinear" inputs. It has been shown that the presented VSC controller can drive the trajectories of the uncertain systems with "nonlinear" inputs onto the sliding mode. Furthermore, it has been proven that the uncertain systems with "nonlinear" inputs also possess the advantage of insensitivity to the uncertainties and/or disturbances as those systems with "linear" inputs in the sliding mode. Therefore, the investigated VSC method of the current

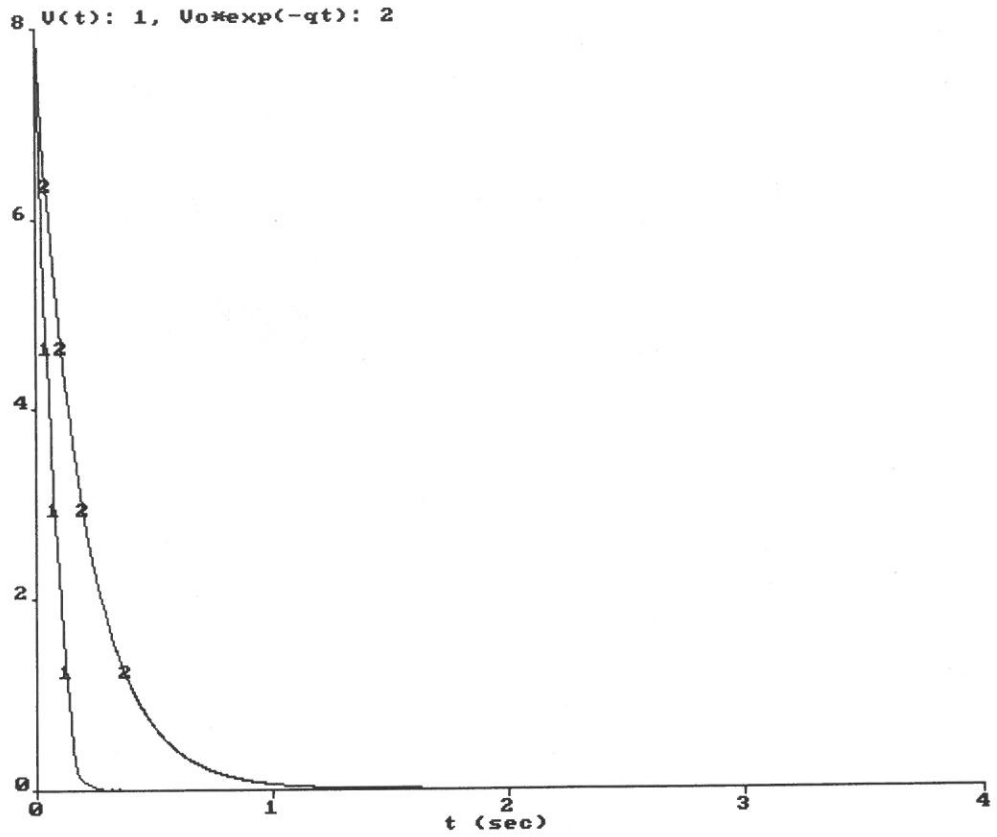



**Fig. 6. The corresponding Lyapunov function $V(t)$ and specified converging speed $V(0)e^{-qt}$.**

work can be applied to more practical physical systems with and/or without input nonlinearities. The converging rate can also be ensured within a specified range.

## ACKNOWLEDGEMENT

The author would like to express his gratitude to the SVD section, Fu Jen Catholic University for the financial support to this research project.

## REFERENCES

(1) F. Jabbari and W. E. Schmitendorf, "A noniterative method for the design of

linear robust controllers," *IEEE Trans. Autom. Control*, Vol. 35, pp. 954-957, (1990).

(2) B. R. Barmish, "Necessary and sufficient conditions for quadratic stabilization of an uncertain system," *J. of Optimization theory and Applications*, Vol. 46, pp. 399-408, (1985).

(3) B. R. Barmish, " Stabilization of uncertain systems via linear control," *IEEE Trans. Autom. Control*, Vol. 28, pp. 848-850, (1983).

(4) I. R. Petersen, "A stabilization algorithm for a class of uncertain linear systems," *System Control Letters*, Vol. 8, pp. 351-357, (1987).

(5) P. P. Khargonekar, I. R. Petersen, and K. Zhou, "Robust stabilization of uncertain linear systems: quadratic stabilizability and $H^{\infty}$ *control theory*," *IEEE Trans. Autom. Control*, Vol. 35, pp. 356-361, (1990).

(6) Y. H. Chen, "Adaptive robust control of uncertain systems with measurement noise," *Automatica*, Vol. 28, pp. 715-728, (1992).

(7) K. Gu and Y. H. Chen, "Linear control quaranteeing stability of uncertain systems via orthogonal decomposition," *Automatica*, Vol. 27, pp. 873-876, (1991).

(8) Z. Qu, "Asymptotic stability of controlling uncertain dynamical systems," *Int. J. Control*, Vol. 59, pp. 1345-1355, (1991).

(9) Y. Ohta, I. R. Gurov, and H. Haneda, "Quadratic stabilization of uncertain Lur'e-Postnikov systems via linear state feedback," *Control Theory and Advanced Technology*, Vol. 8, pp. 353-360, (1992).

(10) D. Recker and P. V. Kokotovic, "Indirect adaptive nonlinear control of discrete-time systems containing a deadzone," *IEEE Proc. of the 32nd Conf. on Decision and Control*, San Antonio, TX, pp. 2647-2653, (1993).

(11) G. Tao and P. V. Kokotovic, "Adaptive control of continuous-time systems with unknown backlash," *IEEE Trans. on Automatic Control*, Vol. 40, pp. 1083-1087, (1995).

(12) N. Chalhoub and X. Zhang, " Modeling and control of backlash in the drive mechanism of a radially rotating compliant beam," *ASME J. of Dynamic Systems*, Vol. 118, pp. 158-161, (1996).

(13) W. M. Haddad and V. Kapila, " Antiwindup controllers for systems with input nonlinearities," *J. of Guidance, Control, and Dynamics*, Vol. 19,

pp. 1387-1390, (1996).

(14) M. Vidyasagar, *Nonlinear Systems Analysis*, Prentice Hall, Englewood Cliff, NJ, (1993).

(15) J. J. Slotine, *Applied Nonlinear Control*, Prentice Hall, Englewood Cliff, NJ, (1991).

(16) U. Itkis, *Control System of Variable Structure*, Wiley, New York, (1976).

(17) V. I. Utkin, *Sliding Mode and Their Applications in Variable Structure Systems*, MIR Editors, Moscow, (1978).

(18) F. Chang, S. Twu, and S. Chang, "Adaptive chattering alleviation of variable structure system control," *IEE Proceeding part D*, Vol. 137, pp. 31-39, (1990).

(19) C. Dorling and A. Zinober, "Robust hyperplane design in multivariable variable structure control systems," *Int. J. Control*, Vol. 48, pp. 2043-2054, (1988).

(20) N. Gough, Z. Ismail, and R. King, "Analysis of variable structure system with sliding modes," *Int. J. System Sciences*, Vol. 15, pp. 401-409, (1984).

(21) B. White and P. Silson, "Reachability in variable structure control systems," *IEE Proceeding part D*, Vol. 131, pp. 85-91, (1984).

(22) X. Xu, Y. Wu, and W. Huang, "Variable-structure control of decentralized model-reference adaptive systems, *IEE Proceeding part D*, Vol. 137, pp. 302-306, (1990).

(23) K. K. Shyu and J. J. Yan, "Variable-structure model following adaptive control for systems with time-varying delay," *Control-Theory and Advanced Technology*, Vol. 10, pp. 513-521, (1994).

(24) K. K. Shyu and C. Y. Liu, "Variable structure controller design for robust tracking and model following," *J. of Guidance, Control, and Dynamics*, Vol. 19, pp. 1395-1397, (1996).

(25) B. Drazenovic, "The invariance condition in variable structure systems," *Automatica*, Vol. 5, pp. 287-295, (1969).

86年10月21日 收稿
86年12月15日 接受

# 可變結構控制於不確定 Lur'e-Postnikov 系統之應用

徐 國 政
輔仁大學電子工程學系

## 摘 要

本文針對不確定 Lur'e-Postnikov 系統，提出一強健控制方法。研究指出：在滑動模式中，具有非線性輸入之系統，仍然擁有線性輸入系統其不受不確定因子或外界干擾影響的特性。故所發展之可變結構控制器，能使不確定 Lur'e-Postnikov 系統的軌跡，進入所預定之滑動模式；然而控制計所需之資訊，僅爲不確定因子與非線性輸入函數之邊界。同時，此滑動模式之收斂速度，可被限制在一預定的指數衰減範圍之內。

**關鍵詞：** Lur'e-Postnikov 系統，可變結構控制，強健控制，非線性輸入，連續非線性。

# 半成批電源管理方法應用在掌上型多媒體手機

白 英 文
輔仁大學電子工程學系

## 摘　　要

電源管理技術是行動計算建置所需關鍵技術之一，尤其是針對本文所界定之掌上型多媒體手機。由於網路傳輸的變異性，手機經常處於閒置狀態等待多媒體資訊抵達。閒置狀態下的手機浪費了寶貴電源，造成必需經常充電。本文提出一種半成批電源管理方法將閒置手機切換狀態至低功率狀態，直到手機累積到足夠多之多媒體資訊，再將手機狀態切換至正常處理及顯示狀態，從機率統計模式及模擬結果顯示，半成批電源方法可以節省電源消耗。爲簡化分析我們採用 Possion 分佈模型且將手機操作區分爲低功率狀態、轉換狀態以及正常操作狀態。

**關鍵詞：** 掌上型多媒體、電源管理、半成批、佇列、延遲變異

## 一、導　　論

行動計算系統所需要的電源管理技術最近有一些研究正在進行〔1，2，3〕。電源管理的目的在於節省耗電而能延長行動計算手機操作時間而降低再重複充電的頻繁度。基本上，透過手機狀況的偵測，判定機器爲閒置時，系統將關掉部份子系統以節約用電〔4，5〕。一般而言，行動計算手機系統內有計時器用以計數閒置區間再用以決策進而切換系統進入睡眠狀態，也就是低耗電狀態。省電的原因在於統計上手機處於“閒置”的時間比“忙碌”的時間長。在節電最差的情形下，手機頻繁切換在“忙碌”與“閒置”之間，手機操作性能會下降且其耗電由於切換狀態可能輕微增加。

爲了處理手機“忙碌”與“閒置”動態間插問題，我們提出半成批電源管理方法，尤其針對行動計算中之掌上型多媒體手機由於多媒體網路資訊流量的變異性（variation）容易造成手機“忙碌”與“閒置”切換頻繁情形。如果採用半成批方法，手機可以累積抵達的多媒體資訊。直到存在佇列（queue）內的存量達到設定的臨界時手機從“閒置狀態”再切換至“忙碌狀態”，進行處理所收到的多媒體資訊，讓使用者看到一般足夠長的多媒體資訊，再將手機狀態切換至閒置，進行下一階段多媒體資訊累積，本文所指“閒置”是除了命令佇列及抵達資訊佇列工作之外，其餘子系統可視其需求關掉其供電或降低其供電用以節省耗電。由於掌上型多媒體手機有其應用特質，故首先，界定其規格〔6，7，8〕，並指出適用於半成批電源管理方法之所在。接著我們採用狀態轉移模型代表行動使用者操作特性。跟著以狀態轉移模型爲基礎，加以多媒體資訊佇列形成半成批電源管理方法，並且推估其節省效率，最後總結論點並提出進一步研究方向：可調式成批電源管理方法或適應式成批電源管理方法可能進一步節省多媒體手機耗電。

## 二、掌上型多媒體手機規格界定

圖 1 顯示多媒體資訊系統架構，行動使用者手機端透過無線數據傳輸或紅外線獲得網路多媒體資訊；使用者亦可透過插卡片（如 PCMCIA CARD）獲得局部封閉性多媒體資訊（如電子書）。

圖 2 顯示各類計算機網路特性，包括其資訊傳輸速率、延遲。一般而言，涵蓋範圍越小之單質性網路，其資訊傳輸速率愈快（如：100Mbps 以上），也比較能提供動態影像之傳輸。相對而言涵蓋範圍愈廣之網路，其資訊傳輸速率愈慢（如：64kbps），比較適合提供文字型資訊或低解析度多媒體畫面，因爲異質性網路，軟硬體介面轉換，長途轉接傳輸，其資訊流量自然遞減。

從圖 2 計算機網路傳輸延遲（transmission delay）變異性有其隨機性，尤其在各類網路連結之後，其總延遲變異性（delay variance）爲各類延遲變異性之總和〔15，16〕，也因此手機接受多媒體資訊斷斷續續的特性是經常存在的，如果手機一直處於正常操作耗電狀態，則由於長時間等待資訊抵達，浪費電源而需要經常再充電，對行動使用者將造成操作上的不方便〔14〕。所以以累積網路抵達資訊而形成半成批處理的電源管理方法有其節省耗電的潛力。

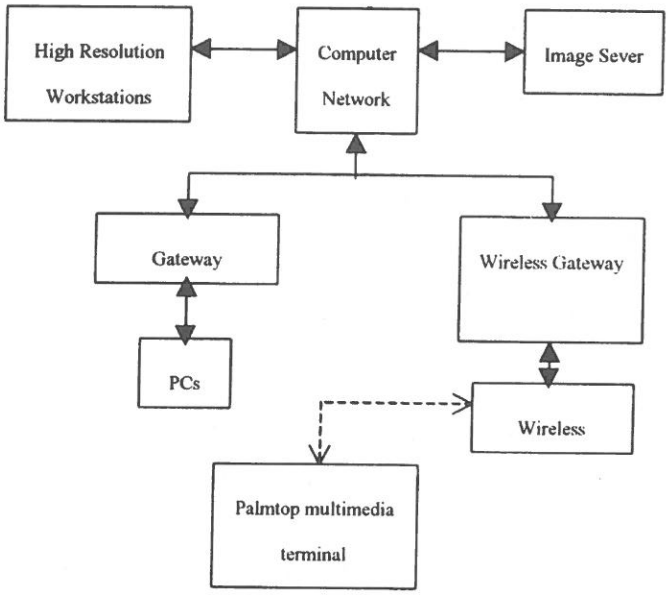



**圖 1　多媒體資訊系統架構〔16〕**

| Network | Bandwidth (Mbps) | Dedicated or Shared | Transmission Delay | Delay Variance |
|---|---|---|---|---|
| Ethernet | 10 | Shared | Random | ∞ |
| Iso-Ethernet (isochronous part) | 10+6 | Shared | Fixed<1ms | 0 |
| Token Ring | 4/16 | Shared | Configuration dependent<20ms | Max |
| 100Base-T | 100 | Shared | Random | ∞ |
| Demand Priority | 100 | Shared | Configuration dependent<10ms | Max |
| FDDI | 2×100 | Shared | Configuration dependent | Max |
| FDDI Ⅱ (isochronous part) | n×6 | Shared | Fixed<1ms | 0 |
| DQDB QA Access | 2×100 | Shared | Random | ∞ |
| DQDB PA Access | 2×100 | Shared | Fixed | 0 |
| X.25 | <2 | Dedicated | Random | ∞ |
| Frame Relay | <50 | Dedicated | Random | ∞ |
| ISDN | n×0.064 | Dedicated | Fixed<10ms | 0 |
| ATM | | Dedicated | Bounded<10ms | Max (AAL 5) |

∞ = asynchronous network without delay jitter control

max = synchronous network with delay variance between 0 and max delay

0 = isochronous network with constant delay

**圖 2　計算機網路傳輸延遲特性**

圖 3 顯示多媒體資訊品質與傳輸頻寬之關係〔12〕，其中文章資訊、靜態影像、聲音（CD）、動畫、動態影像所需傳輸頻寬（非壓縮）分別爲 16kbps、600kbps、704kbps、2.4Mbps、27.7Mbps，其傳輸頻寬有其上下界限，也是產生傳輸延遲變異之原因。

目前（1997/10）手持式小電腦特性，爲了便於攜帶操作，一般重量低於 1.5 磅，體積小於 $1.2\times7\times3.6$ 英吋$^3$（in$^3$）。人機介面採用點觸螢幕、筆勢及手寫辨識、螢幕內建軟體鍵盤、或迷你鍵盤。顯示螢幕大約爲 $5in\times3in$，解析度爲 $480\times320$ 圖素（pixels），符合 64k Video phone（數百 kpixels/frame），由於手機體積重量小，所能儲存電源有限，所以用電效率則更顯得重要。

未來掌上型多媒體手機在規格上略小於手持式小電腦，更易於攜帶操作。人機介面「劇情式整體資訊服務方式」，透過點觸、筆勢、語音等推動。可攜式多媒體服務採用此類人機介面，原因是因爲在手上操作環境下，無法從事太複雜之操作。針對行動操作環境及節電機制，我們建議掌上型多媒體手機內配備兩個佇列：一個命令佇列，用以暫存使用者命令，其操作特性在第三節討論。另一個爲多媒體資訊佇列，其半成批節電管理方法在第四節討論。

## 三、行動使用者操作模式

理想上，一般行動使用者對掌上型多媒體手機下達命令後，期望馬上獲得想要之服務，如果手機或網路反應太慢，可能讓使用者誤會手機未曾接受命令，習慣上，使用者會繼續下達命令，因而讓機器忙於服務使用者，但是使用者並不了解機器眼前畫面的資訊是針對那一個命令的結果。使用者必須重新操作，造成挫折感。故我們建議，在手機接受命令以後，人機介面應在畫面顯示手機現在工作狀況，減少使用者誤會。

行動使用者操作模式因人、事、物等環境而異，並且與個人習慣相關，所以其操作模型極爲複雜。爲了簡化操作模型，我們採用機率統計的觀點，引用狀態轉移模型（State-transition model），如圖 4 所示。

$\lambda_0$，$\lambda_1$，…，$\lambda_{N-1}$使用者下命令速度，$\mu_1$，$\mu_2$，…，$\mu_N$ 手機服務命令速度，狀態 0 爲手機閒置狀態，狀態 1 至 $N$ 爲手機忙碌狀態，假設使用者僅能極短時間內下 $N$ 個命令，也就是留在佇列（queue）內部僅能有 $N$ 個命令，超過的命令將被手機遺失。

假設命令抵達之機率爲柏松分布（Possion distribution），以 $M/M/1$ 佇列簡略

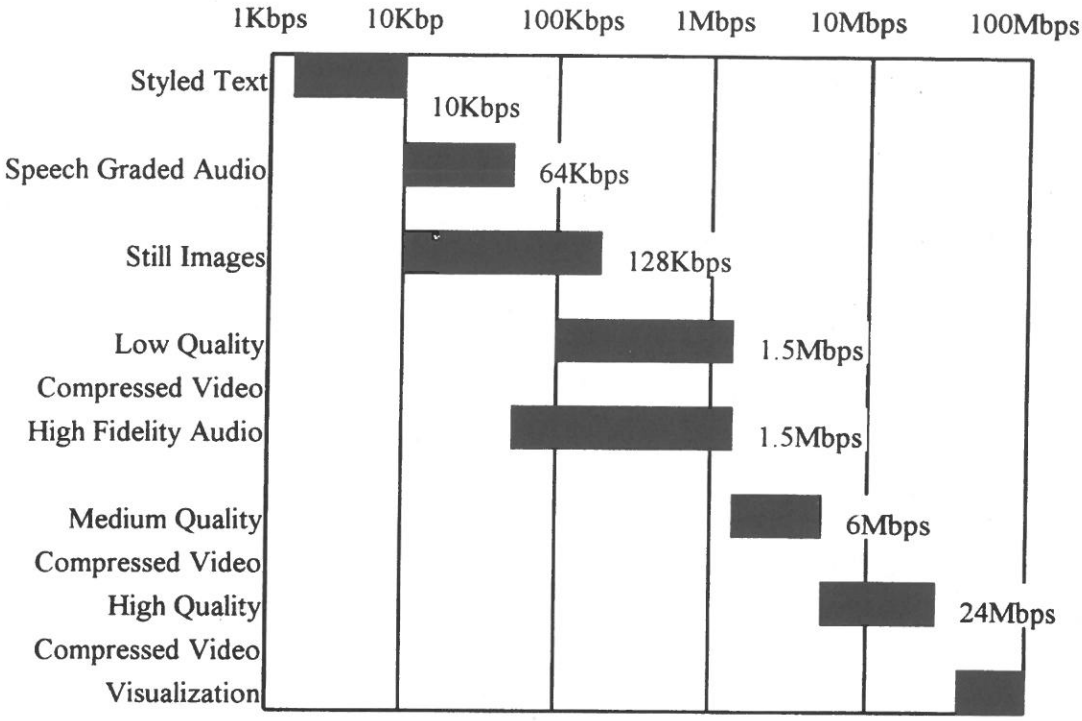



**圖 3　多媒體物件品質與傳輸頻寬〔16〕**

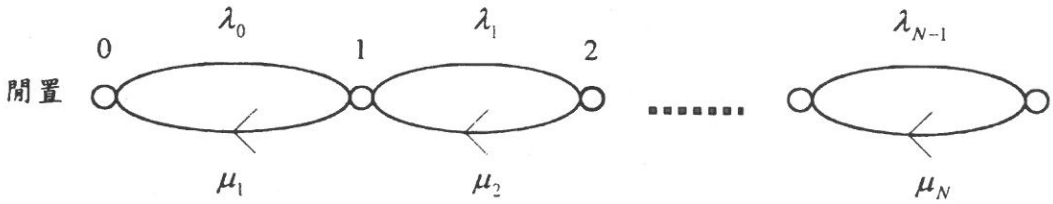



**圖 4　狀態轉移模型**

分析〔**16**〕，其中在狀態轉移模型中各個狀態存在的機率，提供操作者模式的概略情形，基於狀態平衡觀念，手機狀態機率分析如下：

$$\mu_1 P_1 = \lambda_0 P_0, P_1 = \frac{\lambda_0}{\mu_1} P_0 \tag{1}$$

當 $N \geqq n \geqq 1$，其通式爲

$$(\lambda_n + \mu_n)P_n = \mu_{n+1}P_{n+1} + \lambda_{n-1}P_{n-1} \tag{2}$$

利用遞迴疊代法，我們獲得

$$P_n = \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}P_0 \tag{3}$$

因爲命令佇列長度爲 $N$，所有狀態機率總和爲 1：

$$\sum_{n=0}^{N} P_n + \sum_{n=n+1}^{\infty} P_n = 1$$

假設 $\lambda_0 = \lambda_1 = \cdots = \lambda_{n-1} = \lambda$，$\mu_1 = \mu_2 = \cdots = \mu_n = \mu$：

①則可估算機器閒置機率 $P_0$：

$$P_0 = \frac{(1-\rho)}{1-\rho^{N+1}}, \rho = \frac{\lambda}{\mu} < 1 \tag{4}$$

②而機器忙碌機率爲 $1 - P_0$

③使用者命令被遺棄機率爲：

$$P_N = \frac{(1-\rho)\rho^N}{1-\rho^{N+1}} \tag{5}$$

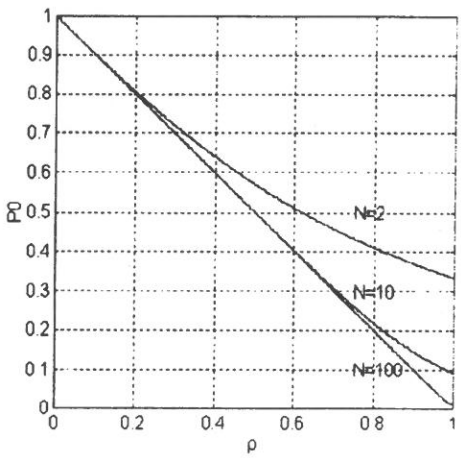



**圖 5　手機閒置機率**

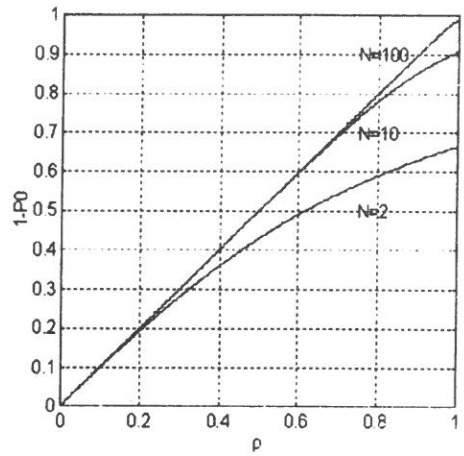



**圖 6　手機忙碌機率 1－P0**

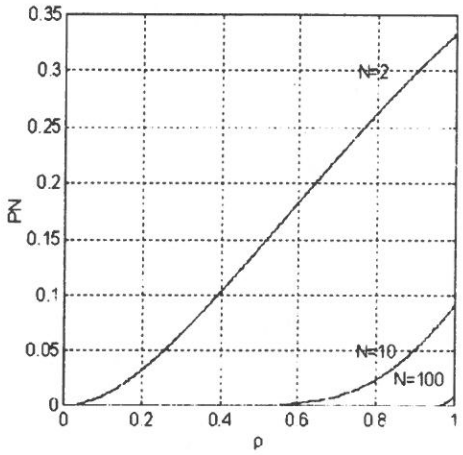



**圖 7　使用者命令被遺棄機率**

圖 5、6、7 分別表示 $P_0$、$1-P_0$ 和與 ρ 之關係，一般而言，手機對命令的服務速率 $\mu$ 應大大於命令抵達速率 λ，也就是 $\rho << 1$，再這種情形下：

$$P_0 = \frac{(1-\rho)}{1-\rho^{N+1}} \approx 1, \rho << 1 \qquad (6)$$

$P_0$ 近於 1，代表手機服務速率大大於命令抵達速率時，手機幾乎在閒置狀態。

而 $P_0 \approx 1$，所以 $1 - P_0 \approx 0$，$1 - P_0 \approx 0$，代表手機在忙碌機率爲 0，至於命令被遺棄機率 $P_N$：

$$P_N = \frac{(1-\rho)\rho^N}{1-\rho^{N+1}} \approx 0, \rho^N \approx 0 \tag{7}$$

## 四、半成批電源管理方法評估

半成批資訊處理系統如圖 8 所示，通訊服務子系統內佇列累積至足夠多之資訊，本系統之多媒體資訊處理子系統和資訊顯示子系統才被啓動處理所接受之資訊，也因此在等待資訊狀態下，本系統中有兩個子系統是處於閒置狀態，可以將它們切換至低耗電狀態，所以整體而言其平均耗電可降低。

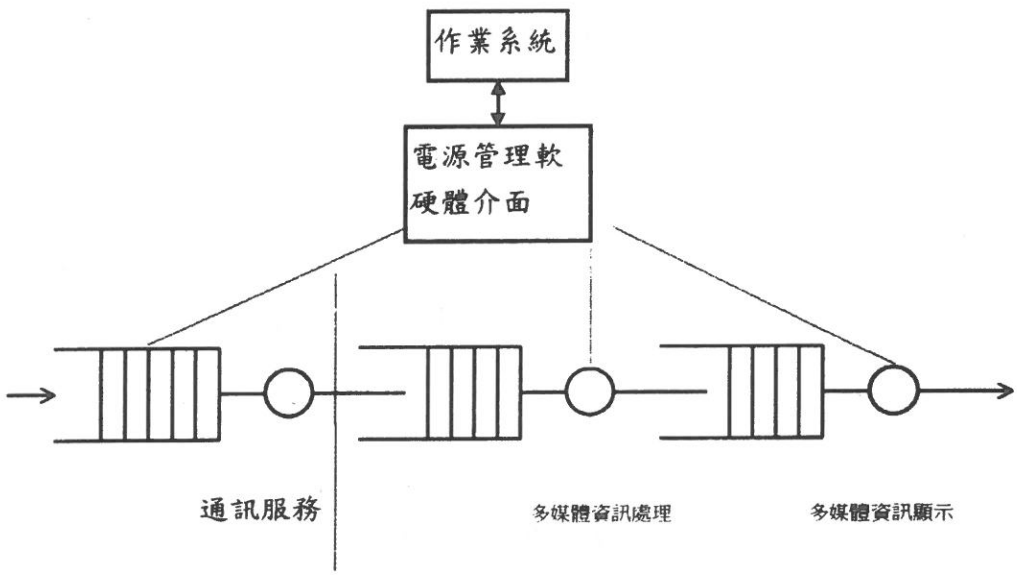



**圖 8　半成批資訊處理系統**

實際手機操作是受限於資訊抵達之隨機特性，也就是網路傳輸特性及狀況決定資訊抵達之特性。故其精確推估有其困難，所以我們採用常用機率分佈從事近似之推估。基本上，沿用第四節之狀態轉移模型，將使用者命令佇列改成多媒體資訊佇列，透過佇列長度之設定與佇列利用度（ρ）之調整，近似推估手機系統等待資訊之

閒置機率以及節電效率。經由穩態 $M/M/1$ 佇列狀態轉移機率分析〔10，11〕，從公式（3）我們知道

$$P_n = \rho^n P_0, P_0 = 1 - \rho, \rho = \frac{\lambda}{\mu} \tag{8}$$

λ 爲平均資訊抵達速度，$\mu$ 爲平均資訊服務速度，$Pn$ 爲資訊佇列內存有 $n$ 筆資訊的機率。

如果佇列長度 L 爲系統處於等待之機率 Pwaiting

$$\begin{aligned} P_{waiting} &= P_0 + P_1 + \cdots + P_L = \sum_{n=0}^{L} \rho^n P_0 \\ &= P_0 + \rho P_0 + \rho^2 P_0 + \cdots + \rho^L P_0 = 1 - \rho^{L+1} \end{aligned} \tag{9}$$

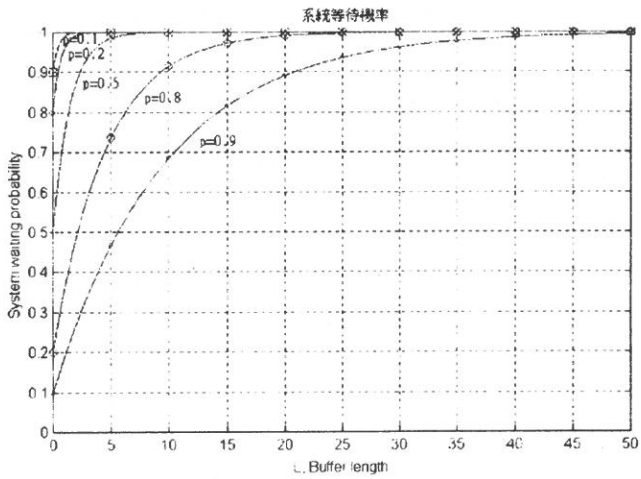



**圖 9　系統等待機率**

圖 9 顯示多媒體資訊佇列長度（L）設定愈長之系統，其系統處於等待之機率愈高，而資訊處理速率比抵達速率愈高之系統，其系統處於等待之機率亦高。

至於系統耗電情形則依系統所處狀態而定，其系統狀態可區分爲等待狀態、等待和正常切換狀態、正常狀態，三種狀態耗電分別爲 $xPOWER$、$yPOWER$ 和 $POWER$，故總功率可爲 $POWER_{total}$。

$$POWER_{total} = xPOWER \cdot P_{\substack{waiting \\ states}} + yPOWER \cdot P_{\substack{transistion \\ states}} + POWER \cdot P_{\substack{normal\ operation \\ states}}$$

(10)

$$P_{\substack{waiting \\ states}} = 1 - \rho^{n+1} \tag{11}$$

$$P_{transistion} = \rho^{n}(1-\rho) \tag{12}$$

$$P_{normal\ operation} = \rho^{n+1} \tag{13}$$

從多個實用系統統計〔3，4〕，我們知道

$$x \cong 0.25$$

$$y \cong 1.25 \tag{14}$$

x，y 由實際系統推估取用，而將 POWER 定爲 100%。

圖 10 顯示方程式（10）之結果。佇列長度（n）愈長，系統等待機率較大，多數子系統處於低功率狀態，故其耗電低於 100%。另外，當資訊抵達速率小小於服務速率時，系統等待機率較大，其耗電也低於 100%。

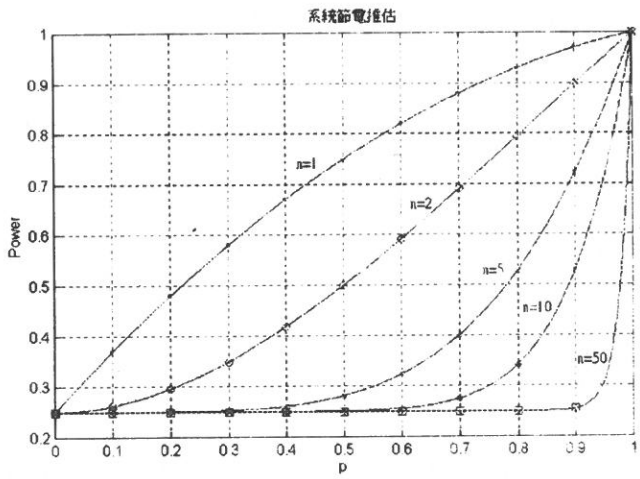



**圖 10　系統節電推估**

## 五、結論與討論

掌上型多媒體手機提供即時性資訊服務，是未來工商社會提高競爭力的利器，

預計在公元 2000 年市場規模達 500 萬台以上。由於隨身攜帶，故其重量應低於 1．5 磅以下，體積易於手持，人機介面以點觸圖型視窗爲主，輔以筆勢、筆式辨認，以及特定簡易詞句語音命令辨識。由於在手機重量限制下，其電池容量不可能太大，所以電源管理變成非常重要。

電源軟體管理模組，對於沒有動作的子系統採取特定的控制措施，用以節省功率消耗。典型的措施包括：停止供電，減慢時鐘訊號速度或停止時鐘驅動訊號以達到省電的目的〔9〕。

行動使用者操作模式因人、事、物等環境因素而異。並且與個人操作習慣相關，所以其操作模型極爲複雜，爲了簡化操作模式，我們採用機率統計的觀點，引用狀態轉移模型，粗略估算手機閒置機率、忙碌機率，以及使用者命令被遺棄機率。

半成批電源管理方法主要著眼點在通訊子系統內佇列累積足夠多的資訊，再進行處理和顯示，在等待資訊期間處理及顯示子系統被切換至低耗電狀態，因而降低整體耗電，降低手機再充電之頻繁度。

多媒體網路資訊抵達手機情形具有隨機特性，前後筆資訊有其獨特關聯性，未來研究將朝向此類資訊關聯性與節電方法之研究。

## 六、謝　　誌

感謝聖言會對本研究的補助。

## 參 考 文 獻

(1) Z. J. lemnios and K. J. Gabriel, "Low-Power Electronics," IEEE Design & Test of Computer, Vol. 11, No. 4, PP. 8-13, Winter (1994).

(2) A. Chandrakasan, A. Burstein, and R. W. Brodersen, "A Low Power Chipset for Portable Multimedia Applications," 1994 IEEE International Soild-State Circuit Conference, PP. 82-83, (1994).

(3) T. H. Meng, B. M. gordon, E. K. Tsern, and A. C. Hung, "Portable Video-on-Demand in Wireless Communication," Processings of the IEEE, Vol. 83, No. 4, PP. 659-681, April (1995).

(4) J. M. C. Stork, "Technology Leverage for Ultra-Low Power Information Systems," Processings of the IEEE, Vol. 83, No. 4, PP. 607-61, April

(1995).

(5) E. P. Harris, S. W. Depp, W. E. Pence, S. Kirkpatrick, M. Sri-jayantha, and R. R. Troutman, "Technology Directions for Portable Computers," Processings of the IEEE Vol. 83, No. 4, PP. 636-658, April (1995).

(6) A. A. Abidi, "Low-Power Radio-Frequency IC' s for Protable Communication," Processings of the IEEE Vol. 83, No. 4, PP. 544-569, April (1995).

(7) C. A. Waraick, "Trends and Limits in the" Talk Time "of Personal communicatiors," Proceedings of the IEEE, Vol. 83, No. 4, April (1995).

(8) J. D. Meindl, "Low Power Microelectronics: Retrospect and Prospect," Proceedings of the IEEE, Vol. 83, No. 4, April (1995).

(9) F. N. Nojm, "A Survey of Power Estimation Techniques in VLSI Circuits," IEEE Trans. VLSI System, Vol. 2, No. 4, PP. 446-455, DEC. (1994).

(10) C. L. Su, C. Y. Tsui, and A. M. Despain, "Saving Power in the Control Path of Embedded Processors," IEEE Design & Test of Computer, PP. 24-30, Winter (1994).

(11) S. A. Khan, V. K. Madisetti, "System Partitioning of MCM for Low Power," PP. 41-52, Spring (1995).

(12) R. Baldazo, "Portable Multimedia," Byte, Vol. 20, No. 6, PP. 169 (1), June (1995).

(13) W. Wong, "Power Management Gets Smart," LAN Magazine, Vol. 10, No. 4, PP. 153 (4), April (1995).

(14) M. R. Zimmerman, "Phoenix Introduces Customizable BIOS; Improves Power Managment," PC Week Vol. 12, No. 3, PP. 41 (2) Jan 23, (1995).

(15) K. Pahlavan and A. H. Levesque, Wireless Information Networks, John Wiley & Sons, Inc. (1995).

(16) B. O. Szuprowicz, Multimedia Networking and Communications, Computer Technology Research Corp. (1994).

86年10月21日 收稿
86年12月 8日 修正
86年12月24日 接受

# Semi-batch Power Management Methods for a Palmtop Multimedia Terminal

YING-WEN BAI
*Department of Electronic Engineering*
*Fu-Jen University*
*Taipei, Taiwan 242, R.O.C.*

## ABSTRACT

Power management technology is one of enable technologys for the implementation of mobile computing. Especially, for a palmtop multimedia terminal, due to the variance of the network transmission speed, the terminal may wait the multimedia data coming and do noting except wasting power. To redue power consumption, this paper propose a semi-batch power management method which can switch the handheld machine to waiting state with low power consumption. Until the terminal accumulate certain amount of multimedia information, it can be switched to normal state with normal power consumption in order to process and display the block of information. From mathematical point of view and simulation result, we show the semi-batch power management method can save power consumption based on the combination of low power state, transistion state, and normal power state. For the simplication of the combination modeling, we use Possion distribution model for the analysis.

***Key Words:*** Palmtop multimedia, Power management, Semi-batch, Queue, Delay variance.

# A Note on Estimation Algebras with Maximal Rank

WEN-LIN CHIOU[1] AND CHENBING LU

*Department of Mathematics,*

*Fu-Jen University*

*Taipei, Taiwan 242, R.O.C.*

## ABSTRACT

The idea of using estimation algebra to construct the finite nonlinear filter was first proposed by Brockett and Clark, Brockett, and Mitter independently. The concept of estimation algebras was proven to be an invaluable tool in the study of nonlinear filtering problems. In his famous talk at the International Congress of Mathematicians in 1983, Brockett proposed to classify all finite dimensional estimation algebras.

In the paper of Chen-Yau-Leung, they introduced a matrix's equation and use it to obtain a classification theorem of finite dimentional esitimation algebras of maximal rank with the state space dimension 4.

In this paper, we provide a different proof concerning the equation of matrix when the state space dimension is less than or equal to 4. Also for arbitrary finite-dimensional state apace, we provide a simple sufficient and necessary condition for a structure result.

***Key Words:*** Nonlinear filter, Estimation algebra.

## INTRODUCTION

The idea of using estimation algebras to construct finite dimensional nonlinear filters was first proposed in Brockett and Clark[1], Brockett[3]and Mitter[12]. The

[1] E-mail:math 1014@fujens.fju.edu.tw. funded by NSC grant NSC-86-2115-M-030-002

concept of estimation algebras was proven to be an invaluable tool in the study of nonlinear filtering problems. In his famous talk at the International Congress of Mathematicians in 1983, Brockett [2] proposed to classify all finite dimensional estimation algebras. Let $n$ be the dimension of the state space. It turns out that all nontrivial finite dimensional estimation algebras are automatically exact with maximal rank if $n = 1$. It follows from the works of Ocone [13], Tam-Wong-Yau [14], and Dong-Tam-Wong-Yau [11] that the finite dimensional estimation algebras are completely classified if $n = 1$. In fact, Dong, Tam, Wong, and Yau have classified all finite dimensional exact estimation algebras with maximal rank of arbitrary finite state space dimension for filtering system (4). For arbitrary finite dimensional state space, under the condition that the drift term is a linear vector field plus a gradient vector field, Yau [15] have classified all finite dimensional estimation algebras with maximal rank of filtering system (4). Chiou-Yau [5], Chen-Leung-Yau [6] and Chen-Yau-Leung [7] have classified all finite dimensional estimation algebras with maximal rank of filtering system (4) for $n = 1,2$, $n = 3$, and $n = 4$, respectively. Chiou [8] consider a different filtering system (cf. system equation (1) with $g(x(t) =$ nonsingular constant matrix). The author obtain a similar classiification theorem as in Yau [15], and have classified all finite dimensional estimation algebras with maximal rank for $n \leq 4$ in this filtering system.

In the paper of Chen-Yau-Leung [7], they obtain a classification theorem offinite dimentional esitimation algebras of maximal rank with the state space dimension 4. The key point of the above classification theorem 3.2) showed in Chen-Yau-Leung is to prove that $w_{ij}$ are all constants for $1 \leq i, j \leq n$. And the three authors of the paper introduced a matrix's equation (cf. Theorem 3.1) to deal with the key point of the main theorem.

In this paper, we provide a different proof concerning the equation of matrix when the state space dimension is less than or equal to 4. And for arbitrary finite-dimensional state apace, we provide a sufficient and necessary condition concerning the matrix's equation.

## THE BASIC CONCEPTS

In this section, we will recall some basic concepts and results which we need for

the next section. Consider a filtering problem based on the following signal observation model:

$$\begin{cases} dx(t) = f(x(t))dt + g(x(t))dv(t), & x(0) = x_0 \\ dy(t) = h(x(t))dt + dw(t), & y(0) = 0 \end{cases} \tag{1}$$

in which $x, v, y$ and $w$ are respectively $\mathbf{R}^n$, $\mathbf{R}^p$, $\mathbf{R}^m$ and $\mathbf{R}^m$-valued processes, and $v$ and $w$ have components which are independent, standard Brownian processes. We further assume that $n = p, f, h$ are $c^\infty$ smooth functions, and that $g$ is an $n$ by $n$ $c^\infty$ smooth matrix. we will refer to $x(t)$ as the state of the system at time $t$ and to $y(t)$ as the observation at time $t$ .

Let $\rho(t,x)$ denote the conditional probability density of the state given the observation $\{y(s): 0 \leq s \leq t\}$. It is well known (see [10], for example) that $\rho(t,x)$ is given by normalizing a function $\sigma(t,x)$ , which satisfies the following Duncan-Mortensen-Zakai equation (see [16], for example):

$$d\sigma(t,x) = L_0\sigma(t,x)dt + \sum_{i=1}^{m} L\sigma(t,x)dyi(t), \quad \sigma(0,x) = \sigma_0 \tag{2}$$

Where

$$L_0 = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}(gg^t)_{ij} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_i - \frac{1}{2}\sum_{i=1}^{m} h_i^2$$

and for $i = 1, \cdots, m$ , $L_i$ is the zero degree differential operaror of multiplication by $h_i$. $\sigma_0$ is the probability density of the ini tial point $x_0$ . In this paper, we will assume $\sigma_0$ is a $c^\infty$ function.

Equation (2) is a stochastic partial differential equation. In real applications, we are interesred in constructing state estimators from observed sample paths with some property of robustness, Davis in [10] studied this problem and proposed some robust algorithms.

In our case, his basic idea reduce to defining a new unnormalized density

$$u(t,x) = exp\Big(-\sum_{i=1}^{m} h_i(x)y_i(t)\Big)\sigma(t,x)$$

It is easy to show that $u(t,x)$ satisfies the following time varying partial differential equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = L_0\xi(t,x) + \sum\limits_{i=1}^{m} y_i(t)[L_0, L_i]\xi(t,x) \\ u(0,x) = \sigma_0 \end{cases} \tag{3}$$

Here we have used the following notation.

**Definition.** If X, Y are differential operators. The Lie bracket $[X, Y]$ of X and Y is defined by $[X, Y]\phi = X(Y\phi) - Y(X\phi)$, for any $c^\infty$ function $\phi$

In this paper, we consider filtering systems with $g(x(t))$ is an orthogonal matrix in the filtering equation (1):

$$\begin{cases} dx(t) = f(x(t))dt + g(x(t))dv(t), & x(0) = x_0 \\ dy(t) = h(x(t))dt + dw(t) & y(0) = 0 \\ g(x(t)) \text{ is an orthogonal matrix} \end{cases} \tag{4}$$

Then

$$L_0 = \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^{n} f_i\frac{\partial}{\partial x_i} - \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} - \frac{1}{2}\sum_{i=1}^{m} h_i^2.$$

Let

$$D_i = \frac{\partial}{\partial x_i} - f_i,$$

and

$$\eta = \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{n} h_i^2.$$

Then we have

$$L_0 = \frac{1}{2}\Big(\sum_{i=1}^{n} D_i^2 - \eta\Big)$$

**Property 2.1.**

$(i)[XY, Z] = X[Y, Z] + [X, Z]Y$, where X, Y, Z are differential operators.

$(ii)[gD_i, h] = g\dfrac{\partial h}{\partial x_i}$, $D_i = \dfrac{\partial}{\partial x_i} - f_i$, g and h are smooth functions

$(iii)[gD_i, h] = -gh\omega_{ij} + g\dfrac{\partial h}{\partial x_i}D_j - h\dfrac{\partial g}{\partial x_j}D_i$, where $\omega_{ji} = [D_i, D_j] = \dfrac{\partial f_i}{\partial x_j} - \dfrac{\partial f_j}{\partial x_i}$

$$(iv)[gD_i^2, h] = 2g\frac{\partial h}{\partial x_i}D_i + g\frac{\partial^2 h}{\partial x_i^2}$$

$$(v)[D_i^2, hD_j] = 2\frac{\partial h}{\partial x_i}D_iD_j - 2h\omega_{ij}D_i + \frac{\partial^2 h}{\partial x_i^2}D_j - h\frac{\partial \omega_{ij}}{\partial x_i}$$

$$(vi)[D_i^2, D_j^2] = 4\omega_{ji}D_jD_i + 2\frac{\partial \omega_{ji}}{\partial x_j}D_i + 2\frac{\partial \omega_{ji}}{\partial x_i}D_j + \frac{\partial^2 \omega_{ji}}{\partial x_i \partial x_j} + 2\omega_{ji}^2$$

$$(vii)[D_k^2, hD_iD_j] = 2\frac{\partial h}{\partial x_k}D_kD_iD_j + 2h\omega_{jk}D_iD_k + 2h\omega_{ik}D_kD_j$$
$$+ \frac{\partial^2 h}{\partial x_k^2}D_iD_j + 2h\frac{\partial \omega_{jk}}{\partial x_i}D_k + h\frac{\partial \omega_{jk}}{\partial x_k}D_i$$
$$+ h\frac{\partial \omega_{ik}}{\partial x_k}D_j + h\frac{\partial^2 \omega_{jk}}{\partial x_i \partial x_k}.$$

$$(viii)[D_iD_j, hD_k] = \frac{\partial h}{\partial x_j}D_iD_k + \frac{\partial h}{\partial x_i}D_jD_k + h\omega_{kj}D_i + h\omega_{ki}D_j$$
$$+ \frac{\partial^2 h}{\partial x_i \partial x_j}D_k + h\frac{\partial \omega_{kj}}{\partial x_i}$$

**Definition.** The estimation aalgebra $E$ of a filtering problem (4) is defined to be the Lie algebra generater by

$$\{L_0, L_1, \cdots, L_m\}, or\ E = < L_0, L_1, \cdots, L_m >_{L.A.}$$

$E$ is said to be an estimation algebras pf maximal rank if for any $1 \leq i \leq n$. there exists a constant $c_i$ such that $x_i + c_i$ is in $E$. $E$ is exact if there exists a function $\psi$ such that $f_i = \frac{\partial \psi}{\partial x_i}$ for $1 \leq i \leq n$.

We need the following basic results for later discussion.

**Theorem 2.2.** (Ocone [13]) Let $E$ be a finite dimensional estimation algebra. If a funtion $\xi$ is in $E$ , then $\xi$ is a polynomial of degree at most two.

The following property 2.3 and theorem 2.4 have been proved in Yau [16].

**Property 2.3.** $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = \omega_{ij}$ are constants for all $i$ and $j$ if and only if

$$(f_1, \cdots, f_n) = (l_1, \cdots, l_n) + \left(\frac{\partial \psi}{\partial x_1}, \cdots, \frac{\partial \psi}{\partial x_n}\right)$$

where $l_1, \cdots, l_n$ are polynomials of degree one and $\psi$ is a $c^\infty$ function.

**Theorem 2.4.** Let $F(x_1,\cdots,x_n)$ be a polynomial on $\mathfrak{R}^n$ Suppose that there exists a polynomial path $c:\ \mathfrak{R}\to\mathfrak{R}^n$ such that $\lim_{t\to\infty}\| c(t)\| = \infty$ and $\lim_{t\to\infty}F(c(t)) = -\infty$ Then there are no $c^\infty$ functions $f_1,\cdots,f_n$ on $\mathfrak{R}^n$ satisgying the equations.

$$\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F$$

The following theorem 2.5 proved in Yau [15] plays a fundamental role in the classification of finite dimensional estimation algebras.

**Theorem 2.5.** Let $E$ be a finite dimensional estimation algeba of (4) such that $\omega_{ij} = \frac{\partial f_i}{\partial x_i} - \frac{\partial f_i}{\partial x_j}$ are constant functions. If $E$ is of maximal rank, then $E$ is a real vector space of dimension $2n+2$ with basis given by $1, x_1, x_2, \cdots, x_n, D_1, D_2, \cdots, D_n$ and $L_0$.

The following theorem 2.6 and property 2.7 proved by Chen and Yau [4], [5] is important progress in the program of classification of finite dimension estimation algebras.

Let $Q$ be the space of quadratic forms in $n$ variables,. i. e. real vector space spanned by $x_ix_j, 1\le i\le j\le n$ , Let $X = (x_1,\cdots,x_n)'$ . For any quadratic form $p\in Q$ , there exists a symmetric matrix $A$ such that $p(x) = X'AX$ . The rank of the quadratc form $p$ is denoted by $rk(p)$ and is defined to be the rank of the matrix $A$ . We need the following definition to obtain theorem 2.6 and property 2.7.

**Definition.** A fundamental quadratic form of the estimation algebra $E$ is an element $p_0 \in E\cap Q$ , with the greatest positive rank$i.e.\ rk(p_0)\ge rk(p)$ for any $p\in E\cap Q$. The quadratic rank of the estimation algebra $E$ is defined to be $rk(p_0)$ .

**Theorem 2.6.** Let $E$ be a finite dimensional estimation algebra of maximal rank. Let $k$ be the quadratic rank of $E$ . Then

(1) The observation terms $h_i(x), 1\le i\le m$ are affine polynomials.

(2) (a) $\omega_{ij}$ are constants, for $1\le i\le k$ or $1\le j\le k$

(b) $\omega_{ij}$ are degree one polynomials in $x_{k+1},\cdots,x_n$, for $k+1\le i,j\le n$

(3) $\eta = \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2$ is a polynomial of degree 4. Moreover, the homogeneous polynomial of degree 4 part of $\eta$ depends only on $x_{k+1},\cdots,x_n$ variables.

**Property 2.7.** Let $E$ be a finite dimensional estimation algebra of maximal rank. Let $k$

be the quadratic rank of $E$ . Any homogeneous polynomial of degree 2 in $E$ depends only on $x_1, x_2, \cdots, x_k$.

## A REPROOF AND A SIMPLE SUFFICIENT AND NECESSARY CONDITION FOR A STRUCTURE RESULT

In this section, we provide another proof of the following theorem 3.1, which has been proved by Chen, Yau and Leung [7], and a simple sufficient and necessary condition for a structure result when the state space dimension is arbitrarily finite-dimensional.

**Theorem 3.1.** Suppose that $\eta_4$ is a homogeneous polynomial of degree 4 in $n$ variables. If $n \leq 4$ and $\triangle$ is an antisymmetric matrix with each entry a homogeneous polynomial of degree one such that

$$\triangle\triangle^t = \frac{1}{2}H(\eta_4) \tag{5}$$

where $H(\eta_4) = \left(\frac{\partial^2 \eta_4}{\partial x_i \partial x_j}\right)$ is the Hessian matrix of $\eta_4$ , then $\triangle = 0$

**Proof:** Write

$$\frac{1}{2}H(\eta_4) = \sum_{1\leq i\leq j\leq n} H_{ij}x_i x_j$$

$$\triangle = \sum_{i=1}^{n} A_i x_i \tag{6}$$

Since $\triangle$ is an antisymmetric matrix, we have $A_i = -A_i^t$. By assumption $\triangle\triangle^t = \frac{1}{2}H(\eta_4)$ . we have

$$\begin{aligned}
&(\sum_{i=1}^{n} A_i x_i)(\sum_{i=1}^{n} A_i x_i)^t \\
&= (\sum_{i=1}^{n} A_i x_i)(\sum_{i=1}^{n} A_{ti} x_i) \\
&= \sum_{1\leq i,j\leq n} A_i A_j^t x_i x_j (\sum_{i=1}^{n} A_{ti} x_i) \\
&= \sum_{1\leq i<j\leq n} A_i A_j^t x_i x_j + \sum_{1\leq j<i\leq n} A_i A_j^t x_i x_j + \sum_{i=1}^{n} A_i A_j^t x_i x_i \\
&= \sum_{1\leq i<j\leq n} (A_i(-A_j) + A_j(-A_j))x_i x_j + \sum_{i=1}^{n} A_i(-A_i)x_i x_i
\end{aligned}$$

$$= \sum_{1\le i<j\le n} H_{ij}x_ix_j + \sum_{i=1}^{n} H_{ii}x_ix_i.$$

Hence

$$A_iA_j + A_jA_i = -H_{ij}, A_i^2 = -H_{ii} \tag{7}$$

Let

$$\eta_4 = \sum_{1\le i_1\le i_2\le i_3\le i_4\le n} a(i_1,i_2,i_3,i_4)x_{i_1}x_{i_2}x_{i_3}x_{i_4} \tag{8}$$

where $(i_1,i_2,i_3,i_4)$'s are constants, and let $H_{ij}(i,j)$ denote the $(i,j)$ entry of the matrix $H_{ij}$. Consider the term $a(i,i,j,j)x_i^2x_j^2$ in $\eta_4$ and differentiate it as follows:

$$\frac{\partial^2 a(i,i,j,j)x_i^2x_j^2}{\partial x_i^2} = 2a(i,i,j,j)x_j^2, \frac{\partial^2 a(i,i,j,j)x_i^2x_j^2}{\partial x_j^2} = 2a(i,i,j,j)x_j^2,$$
$$\frac{\partial^2 a(i,i,j,j)x_i^2x_j^2}{\partial x_i\partial x_j} = 4a(i,i,j,j)x_ix_j.$$

Note that

$$H_{jj}(i,i) = 2a(i,i,j,j), H_{ii} = 2a(i,i,j,j), H_{ij}(ij) = 4a(i,i,j,j),$$

so we have

$$H_{jj}(i,i) = H_{ii}(j,j) = \frac{1}{2}H_{ij}(i,j) \tag{9}$$

In view of (7), (8) and (9), we have

$$\sum_i A_j(i,l)A_j(l,i) = \sum_i A_i(j,l)A_i(l,j)$$
$$= \frac{1}{2}\{\sum_l [A_i(i,l)A_j(l,j) + A_j(i,l)A_i(l,j)]\} \tag{10}$$

Since each $A_i$ is an antisymmetric matrix, i. e. $A_i(r,l) = -A_i(l,r)$ for all r, l. We have

$$\sum_i A_j(i,l)^2 = \sum_i A_i(j,l)^2 = \frac{1}{2}\{\sum_l [A_i(i,l)A_j(j,l) + A_j(i,l)A_i(j,l)]\} \tag{11}$$

Similarily, we differentiate the term $a(i,i,i,j)x_i^3x_j$ in $\eta_4$ and get

$$\frac{\partial^2 a(i,i,i,j)x_i^3x_j}{\partial x_i\partial x_j} = 3a(i,i,i,j)x_i^2, \frac{\partial^2 a(i,i,i,j)x_i^3x_j}{\partial x_i^2} = 6a(i,i,i,j)x_ix_j$$

Note that

$$H_{ii}(i,j) = 3a(i,i,i,j), H_{ij}(i,i) = 6a(i,i,i,j).$$

So we have

$$2H_{ii}(i,j) = H_{ij}(i,i). \tag{12}$$

In view of (6) and (12), we have

$$2\{-\sum_l A_i(i,l)A_i(l,j)\} = -\{\sum_l [A_i(i,l)A_j(l,i) + A_j(i,l)A_i(l,i)]\}.$$

Since each $A_i$ is an antisymmetric matrix, i. e. $A_i$ (r, l) $= -A_i$ (l, r) for all r, l, We have

$$\sum_l A_i(i,l)A_i(j,l) = \sum_l [A_i(i,l)A_j(l,i) + \sum_l A_j(i,l)A_i(i,l)],$$

or

$$2\sum_l A_i(i,l)A_i(j,l) = 2\sum_l A_i(i,l)A_j(i,l)$$

So

$$\sum_l A_i(i,l)A_i(j,l) = \sum_l A_i(i,l)A_j(i,l) \tag{13}$$

Let $(\overrightarrow{i,\ j})$ denote the $j$ th column vector of $A_i$, then we can rewrite (12), (13) as follows:

$$|(\overrightarrow{i,j})|^2 = |(\overrightarrow{j,i})|^2 = \frac{1}{2}\{(\overrightarrow{i,j})(\overrightarrow{j,i}) + (\overrightarrow{i,i})(\overrightarrow{j,j})\} \qquad 1 \leq i < j \leq n, \tag{14}$$

$$(\overrightarrow{i,j})(\overrightarrow{j,i}) = (\overrightarrow{i,i})(\overrightarrow{i,j}) \qquad 1 \leq i < j \leq n, \tag{15}$$

We are going to prove $\triangle = 0$, that is to prove each $A_i = 0$ by (6). We only need to prove the state space dimension 4 case. Since the submatrix $((A_i)_{k,l})$, $1 \leq k, l \leq m$ $(m < 4)$ of the matrix $A_i = 0$ if $A_i = 0$.

For state space dimension $n = 4$, since $A_1, A_2, A_3$ and $A_4$ are $4 \times 4$ antisymmetric matrices, we write

$$A_1 = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & A1_{23} & A1_{24} \\ -b & -A1_{23} & 0 & A1_{24} \\ -c & -A1_{24} & - & A1_{34} \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0 & d & A2_{13} & A2_{14} \\ -d & 0 & e & f \\ -A2_{13} & -e & 0 & A2_{34} \\ -A2_{14} & -f & -A2_{34} & 0 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 0 & A3_{12} & g & A3_{14} \\ -A3_{12} & 0 & h & A3_{24} \\ -g & -h & 0 & i \\ -A3_{14} & -A3_{24} & -i & 0 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} 0 & A4_{12} & A4_{13} & j \\ -A4_{12} & 0 & A4_{23} & k \\ -A4_{13} & -A4_{23} & 0 & l \\ -j & -k & -l & 0 \end{pmatrix}$$

By (14), we get thefollowing system of equations

$$|(\overrightarrow{1,2})|^2 = |(\overrightarrow{2,1})|^2 = \frac{1}{2}\{(\overrightarrow{1,2})(\overrightarrow{2,1}) + (\overrightarrow{1,1})(\overrightarrow{2,2})\}$$

$$|(\overrightarrow{1,3})|^2 = |(\overrightarrow{3,1})|^2 = \frac{1}{2}\{(\overrightarrow{1,3})(\overrightarrow{3,1}) + (\overrightarrow{1,1})(\overrightarrow{3,3})\}$$

$$|(\overrightarrow{1,4})|^2 = |(\overrightarrow{4,1})|^2 = \frac{1}{2}\{(\overrightarrow{1,4})(\overrightarrow{4,1}) + (\overrightarrow{1,1})(\overrightarrow{4,4})\}$$

$$|(\overrightarrow{2,3})|^2 = |(\overrightarrow{3,2})|^2 = \frac{1}{2}\{(\overrightarrow{2,3})(\overrightarrow{3,2}) + (\overrightarrow{2,2})(\overrightarrow{3,3})\}$$

$$|(\overrightarrow{2,4})|^2 = |(\overrightarrow{4,2})|^2 = \frac{1}{2}\{(\overrightarrow{2,4})(\overrightarrow{4,2}) + (\overrightarrow{2,2})(\overrightarrow{4,4})\}$$

$$|(\overrightarrow{3,4})|^2 = |(\overrightarrow{4,3})|^2 = \frac{1}{2}\{(\overrightarrow{3,4})(\overrightarrow{4,3}) + (\overrightarrow{3,3})(\overrightarrow{4,4})\} \qquad (16)$$

By (15), we get the following system of equations

$$(\overrightarrow{1,1})(\overrightarrow{2,1}) = (\overrightarrow{1,1})(\overrightarrow{1,2})$$
$$(\overrightarrow{1,1})(\overrightarrow{3,1}) = (\overrightarrow{1,1})(\overrightarrow{1,3})$$
$$(\overrightarrow{1,1})(\overrightarrow{4,1}) = (\overrightarrow{4,4})(\overrightarrow{1,4})$$

$$(\overrightarrow{2,2})(\overrightarrow{3,2}) = (\overrightarrow{2,2})(\overrightarrow{2,3})$$
$$(\overrightarrow{2,2})(\overrightarrow{4,2}) = (\overrightarrow{2,2})(\overrightarrow{2,4})$$
$$(\overrightarrow{3,3})(\overrightarrow{4,3}) = (\overrightarrow{3,3})(\overrightarrow{3,4}) \tag{17}$$

Plug the entries of $A_1, A_2, A_3$ *and* $A_4$ into (16), we have

$$\begin{cases} a^2 + A1_{23}^2 + A1_{24}^2 = d^2 + A2_{13}^2 + A2_{14}^2 \\ \qquad\qquad = \frac{1}{2}(A1_{23}A2_{13} + A1_{24}A2_{14} + be + cf) \\ b^2 + A1_{23}^2 + A1_{34}^2 = A3_{13}^2 + g^2 + A3_{14}^2 \\ \qquad\qquad = \frac{1}{2}(-A1_{23}A3_{12} + A1_{34}A3_{14} - ah + ci) \\ c^2 + A1_{24}^2 + A1_{34}^2 = A4_{12}^2 + A4_{13}^2 + j^2 \\ \qquad\qquad = \frac{1}{2}(-A1_{24}A4_{12} - A1_{34}A4_{13} - ak + bl) \\ A2_{13}^2 + e^2 + A2_{34}^2 = A3_{12}^2 + h^2 + A3_{24}^2 \\ \qquad\qquad = \frac{1}{2}(A2_{13}A3_{12} + A2_{34}A3_{24} + dg + fi) \\ A2_{14}^2 + f^2 + A2_{34}^2 = A4_{12}^2 + A4_{23}^2 + k^2 \\ \qquad\qquad = \frac{1}{2}(A2_{14}A4_{12} - A2_{34}A4_{23} + dj + el) \\ A3_{14}^2 + A3_{24}^2 + i^2 = A4_{13}^2 + A4_{23}^2 + l^2 \\ \qquad\qquad = \frac{1}{2}(A3_{14}A4_{13} + A3_{24}A4_{23} + gj + hk) \end{cases} \tag{18}$$

Expand the system of equation in (18), we have

$$\begin{cases} A1_{23}^2 + A1_{24}^2 - A2_{13}^2 - A2_{14}^2 = d^2 - a^2 & \langle 1 \rangle \\ A1_{23}^2 + A1_{34}^2 - A3_{12}^2 - A3_{14}^2 = g^2 - b^2 & \langle 2 \rangle \\ A1_{24}^2 + A1_{34}^2 - A4_{12}^2 - A4_{13}^2 = j^2 - c^2 & \langle 3 \rangle \\ A2_{13}^2 + A2_{34}^2 - A3_{12}^2 - A3_{24}^2 = h^2 - e^2 & \langle 4 \rangle \\ A2_{14}^2 + A2_{34}^2 - A4_{12}^2 - A4_{23}^2 = k^2 - f^2 & \langle 5 \rangle \\ A3_{14}^2 + A3_{24}^2 - A4_{13}^2 - A4_{23}^2 = l^2 - i^2 & \langle 6 \rangle \end{cases}$$

$$\begin{cases} 2A1_{23}^2 + 2A1_{24}^2 - A1_{23}A2_{13} - A1_{24}A2_{14} = be + cf - 2a^2 & \langle 7\rangle \\ 2A2_{13}^2 + 2A2_{14}^2 - A1_{23}A2_{13} - A1_{24}A2_{14} = be + cf - 2d^2 & \langle 8\rangle \\ 2A1_{23}^2 + 2A1_{34}^2 - A1_{23}A3_{12} - A1_{34}A3_{14} = -ah + ci - 2b^2 & \langle 9\rangle \\ 2A3_{12}^2 + 2A3_{14}^2 + A1_{23}A3_{12} - A1_{34}A3_{14} = -ah + ci - 2g^2 & \langle 10\rangle \\ 2A1_{24}^2 + 2A1_{34}^2 + A1_{24}A4_{12} + A1_{34}A4_{13} = -ak - bl - 2c^2 & \langle 11\rangle \\ 2A4_{12}^2 + 2A4_{13}^2 + A1_{24}A4_{12} + A1_{34}A4_{13} = -ak - bl - 2j^2 & \langle 12\rangle \\ 2A2_{13}^2 + 2A2_{34}^2 - A2_{13}A3_{12} - A2_{34}A3_{24} = +dg + fi - 2e^2 & \langle 13\rangle \\ 2A3_{12}^2 + 2A3_{24}^2 - A2_{13}A3_{12} - A2_{34}A3_{24} = +dg + fi - 2^2 & \langle 14\rangle \\ 2A2_{14}^2 + 2A2_{34}^2 - A2_{14}A4_{12} + A2_{34}A4_{23} = +dg - el - 2f^2 & \langle 15\rangle \\ 2A4_{12}^2 + 2A4_{23}^2 - A2_{14}A4_{12} + A2_{34}A4_{23} = +dg - el - 2f^2 & \langle 16\rangle \\ 2A3_{14}^2 + 2A3_{24}^2 - A3_{14}A4_{13} - A3_{24}A4_{23} = +dg - hk - 2i^2 & \langle 17\rangle \\ 2A4_{13}^2 + 2A4_{23}^2 - A3_{14}A4_{13} - A3_{24}A4_{23} = +dg - hk - 2i^2 & \langle 18\rangle \end{cases}$$

Plug the entries of $A_1, A_2, A_3$ and $A_4$ into (17), and expand it, we have

$$\begin{cases} (b)A2_{13} - (b)A1_{23} + (c)A2_{14} - (c)A1_{24} = -ad & \langle 19\rangle \\ (a)A3_{12} + (a)A1_{23} + (c)A3_{14} - (c)A1_{34} = -bg & \langle 20\rangle \\ (a)A4_{12} + (a)A1_{24} + (b)A3_{13} + (b)A1_{34} = -cj & \langle 21\rangle \\ (d)A3_{12} - (d)A2_{13} + (f)A3_{24} - (f)A2_{34} = -eh & \langle 22\rangle \\ (d)A4_{12} - (d)A2_{14} + (e)A4_{23} - (e)A2_{34} = -fk & \langle 23\rangle \\ (g)A3_{12} - (g)A3_{14} + (h)A4_{23} - (h)A3_{24} = -il & \langle 24\rangle \end{cases}$$

Observe tht eqution $\langle 1\rangle$ is equation $\langle 7\rangle$ plus equation $\langle 8\rangle$,
equation $\langle 2\rangle$ is equation $\langle 9\rangle$ plus equation $\langle 10\rangle$,
equation $\langle 3\rangle$ is equation $\langle 11\rangle$ plus equation $\langle 12\rangle$,
equation $\langle 4\rangle$ is equation $\langle 13\rangle$ plus equation $\langle 14\rangle$,
equation $\langle 5\rangle$ is equation $\langle 15\rangle$ plus equation $\langle 16\rangle$,
and equation $\langle 6\rangle$ is equation $\langle 17\rangle$ plus equation $\langle 18\rangle$.

We are going to prove that the trivial solution is the only solution of the system of equations $\langle 1\rangle$, $\langle 2\rangle$, $\cdots$, $\langle 24\rangle$. That is $(a,b,c,d,e,f,g,h,i,j,k,l) = 0$, and $A1_{23} = A1_{24} = A1_{34} = 0, A2_{13} = A2_{14} = A2_{34} = 0, A3_{12} = A3_{14} = A3_{24} = 0, A4_{12} = A4_{13} = A4_{23} = 0$.

Let equation ⟨25⟩ = ⟨7⟩ + ⟨8⟩ + ⋯ + ⟨18⟩. Then

the left hand side of ⟨25⟩ $= A1_{23}^2 + A2_{13}^2 + (A1_{23} - A2_{13})^2 + A1_{24}^2 + A2_{14}^2 + (A1_{24} - A2_{14})^2 + \cdots$

$A1_{23}^2 + A3_{12}^2 + (A1_{23} + A3_{12})^2 + A1_{34}^2 + A3_{14}^2 + (A1_{34} - A3_{14})^2 + \cdots$

$A1_{24}^2 + A4_{12}^2 + (A1_{24} + A4_{12})^2 + A1_{34}^2 + A4_{13}^2 + (A1_{34} + A4_{14})^2 + \cdots$

$A2_{13}^2 + A3_{12}^2 + (A2_{13} - A3_{12})^2 + A2_{34}^2 + A3_{24}^2 + (A2_{34} - A3_{24})^2 + \cdots$

$A2_{14}^2 + A4_{12}^2 + (A2_{14} - A4_{12})^2 + A2_{34}^2 + A4_{23}^2 + (A2_{34} + A4_{23})^2 + \cdots$

$A3_{14}^2 + A4_{13}^2 + (A3_{14} - A4_{13})^2 + A3_{24}^2 + A4_{23}^2 + (A3_{24} - A4_{23})^2 + \cdots$

$\geq 0$

the right hand side of ⟨25⟩ $= -(b-e)^2 - (c-f)^2 - (a+h)^2 - (c-i)^2$

$-(a+k)^2 - (b+l)^2 - (d-g)^2 - (f-i)^2$

$-(d-j)^2 - (e+l)^2 - (g-j)^2 - (h-k)^2$

$\leq 0$

So we have

the left-hand side of of ⟨25⟩ = the right-hand side of ⟨25⟩ = 0.

Therefore

$A1_{23} = A1_{24} = A2_{13} = A2_{14} = 0$

$A1_{23} = A1_{34} = A3_{12} = A3_{14} = 0$

$A1_{24} = A1_{34} = A4_{12} = A4_{13} = 0$

$A2_{13} = A2_{34} = A3_{12} = A3_{24} = 0$

$A2_{14} = A2_{34} = A4_{12} = A4_{23} = 0$

$A3_{14} = A3_{24} = A4_{13} = A4_{23} = 0$

*and*

$a = -h = -k$

$b = e = -l$

$c = f = i$

$d = g = j$

By equation ⟨1⟩ ⋯ ⟨6⟩, we have $a^2 = b^2 = c^2 = d^2$

Plug the above result into equations ⟨19⟩, ⋯, ⟨24⟩, we have

$$\begin{cases}(b)0-(b)0+(c)0-(c)0=-ad\\(a)0+(a)0+(c)0-(c)0=-bd\\(a)0+(a)0+(b)0+(b)0=-cd\\(d)0-(d)0+(c)0-(c)0=+ab\\(d)0-(d)0+(b)0+(b)0=+ac\\(d)0-(d)0-(a)0+(a)0=+bc\\a^2=b^2=c^2=d^2\end{cases}$$

This implies $a=b=c=d=0$, and $A_1=A_2=A_3=A_4=0$ $Q.E.D.$

The following theorem 3.2 has been proved in Chen-Yau-Leung [7]. For the completeness of this paper, we provide proof here.

**Theorem 3.2.** Suppose that the state space of the filtering system (4) is of dimension $n \leq 4$. If $E$ is the finite dimensional estimation algebra of maximal rank, then the drift term $f$ must be a linear vector field plus a gradient vector field. So $E$ is a real vector space of dimension $2n+2$ with basis given by $1, x_1, x_2, \cdots, x_n, D_1, D_n$ and $L_0$ Moreover $\eta$ is a polynormial of degree at most 2.

**Proof:** Since E is an estimation algebra with maximal rank, there exists constant $c_i$'s such that $x_i + c_i \in E, 1 \leq i \leq n$. By property 2.1

$$[L_0, x_j + c_j] = \frac{1}{2}[\sum_{i=1}^{n} D_i^2 - \eta,\ x_j] = \frac{1}{2}\sum_{i=1}^{n}[D_i^2, x_j] = D_j \in E$$
$$[D_j, x_j + c_j] = 1 \in E$$

Therefore $x_1, \cdots, x_n \in E$. By theorem 2.6, $\omega_{ij}$'s are constants, for $1 \leq i \leq k$ or $1 \leq j \leq k$, where $k$ is the quadratic rank of $E$, and $\omega_{ij}$'s are polynormial of degree one in $x_{k+1}, \cdots, x_n, k+1 \leq i, j \leq n$. Observe that

$$[[L_0, D_j], D_l] = \left[\sum_{i=1}^{n}\left(\omega_{ji}D_i + \frac{1}{2}\frac{\partial \omega_{ji}}{\partial x_i}\right) + \frac{1}{2}\frac{\partial \eta}{\partial x_i}, D_l\right]$$
$$= \sum_{i=1}^{n}\left(\omega_{ji}\omega_{li} - \frac{\partial \omega_{ji}}{\partial x_l}D_i\right) - \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2 \omega_{ji}}{\partial x_l \partial x_i} - \frac{1}{2}\frac{\partial^2 \eta}{\partial x_l \partial x_j}$$

In view of theorem 2. 6 and property 2. 9, we have that

$$\sum_{i=1}^{n}\omega_{ji}\omega_{li} - \frac{1}{2}\frac{\partial^2 \eta}{\partial x_l \partial x_j} \in E, 1 \leq l, j \leq n,$$

and then

$$\sum_{i=k+1}^{n} \beta_{ji}\beta_{li} - \frac{1}{2}\frac{\partial^2 \eta_4}{\partial x_l \partial x_j} \in E, k+1 \leq l, j \leq n.$$

where $\eta_m$ is a homogeneous polynormial of degree m part of $\eta$ (here m = 4) and $\beta_{ij}$ is the homogeneous polymormial of degree 1 part of $\omega_{ij}$. Since $\sum_{i=k+1}^{n} \beta_{ji}\beta_{li} - \frac{1}{2}\frac{\partial^2 \eta_4}{\partial x_1 \partial x_j}$ is a homogeneous polynormial of degree 2 in $E$, it depends only on $x_1, \cdots, x_k$ by property 2.7. On the other hand, it also depends only $x_{k+1}, \cdots, x_n$ by the fact that $\beta_{ji}\beta_{li}$ and $\eta_4$ depends on $x_{k+1}, \cdots, x_n$ and by theorem 2.6. Therefore

$$\sum_{i=k+1}^{n} \beta_{ji}\beta_{li} - \frac{1}{2}\frac{\partial^2 \eta_4}{\partial x_l \partial x_j} = 0, \text{ for } k+1 \leq j, l \leq n.$$

or

$$\triangle \triangle^T = \frac{1}{2} H(\eta_4)$$

where $\triangle = (\beta_{ij}), k+1 \leq i, j \leq n$ is $(n-k) \times (n-k)$ antisymmetric matrix and $H(\eta_4) = \left(\frac{\partial^2 \eta_4}{\partial x_i \partial x_j}\right), k+1 \leq i, j \leq n, H(\eta_4)$ is the Hessian matrix of $\eta_4$. By theorem 2.8, we have $\triangle = 0$. This implies that $\omega_{ij}, 1 \leq i, j \leq n$ are all constants. By property 2.3

$$(f_1, \cdots, f_n) = (l_1, \cdots, l_n) + \left(\frac{\partial \psi}{\partial x_1}, \cdots, \frac{\partial \psi}{\partial x_n}\right)$$

where $l_1, \cdots, l_n$ are linear and $\psi$ is $c^\infty$ smooth. Threrfore by theorem 2.5 $E$ is a real vector space of dimension 2n + 2 with a basis: $\{1, x_1, x_2, \cdots, x_n, D_1, D_2, \cdots, D_n, L_0\}$ Since $\triangle = 0$ implies $\eta_4 = 0$, we have

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{n} h_i^2 = \eta_0 + \eta_1 + \eta_2 + \eta_3$$

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = \eta_0 + \eta_1 + \eta_2 - \sum_{i=1}^{n} h_i^2 + \eta_3$$

If $\eta_3$ were not zero then we can choose a polynormial path $c: \Re \to \Re^n$ such that $\lim_{t\to\infty} \| c(t) \| = \infty$ and $\lim_{t\to\infty} F(c(t)) = -\infty$, where $F = \eta_0 + \eta_1 + \eta_2 + \eta_2 - \sum_{i=1}^{n} h_i^2$. This is impossible by theorem 2.4. So $\eta_3 = 0$ and hence $\eta$ is a polynormial of drgree at most

two. Q.E.D.

We need Lemma A to prove theorem B which is a simple sufficient and necessary condition for a structure result when the state space dimension is arbitrarily finite-dimensional.

**Lemma A.** Let $\vec{a}, \vec{b}$ are vectors in $\Re^n$. Suppose that $|\vec{a}|^2 = |\vec{b}|^2 = \frac{1}{2}\vec{a} \cdot \vec{b}$, then $\vec{a} = \vec{b} = 0$

**Proof:** Suppose $\vec{a} \neq 0$. This implies $\vec{b} \neq 0$. Let $\theta$ be the angle of $\vec{a}$ and $\vec{b}$. Then

$$|\vec{a}|^2 = |\vec{b}|^2 = \frac{1}{2}|\vec{a}|\,|\vec{b}|\cos\theta.$$

So

$$|\vec{a}| = \frac{1}{2}|\vec{b}|\cos\theta, \text{ and } |\vec{b}| = \frac{1}{2}|\vec{a}|\cos\theta$$

Therefore

$$|\vec{a}| = \frac{1}{2}\frac{1}{2}|\vec{a}|\cos\theta^2$$

This implies that $\cos\theta^2 = 4$, which is impossible. So $\vec{a} = \vec{b} = 0$. Q. E. D.

**Property B.** Matrices $A_1 = A_2 = \cdots = A_n = 0$
$\Leftrightarrow (\vec{i,j})(\vec{j,j}) = 0$ for $1 \leq i < j \leq n$

**Proof:** ($\Rightarrow$) Clearly it holds ($\Leftarrow$) By (14)

$$|(\vec{i,j})|^2 = |(\vec{j,i})|^2 = \frac{1}{2}\{(\vec{i,j})(\vec{j,i}) + (\vec{i,i})(\vec{j,j})\}, 1 \leq i < j \leq n \qquad (19)$$

By assumption $(\vec{i,j})(\vec{j,j}) = 0$, and in view of (19). we have

$$|(\vec{i,j})|^2 = |(\vec{j,i})|^2 = \frac{1}{2}\{(\vec{i,j})(\vec{j,i}) \qquad (20)$$

Hence

$$(\vec{i,j}) = (\vec{j,i}) = 0, \qquad 1 \leq i \leq j \leq n$$

we have

$$A_1 = \begin{pmatrix} 0 & A1_{12} & \cdots & A1_{1n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ A2_{21} & 0 & \cdots & A2_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$\vdots$$

$$A_n = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ An_{n1} & An_{n2} \cdots & 0 & \end{pmatrix}$$

Since $A_1, A_2, \cdots, A_n$ are antisymetric matrices, we have. $A_1 = A_2 = \cdots = A_n = 0$

Q.E.D.

## BIBLIOGRAPHY


(1) R. W. Brockett, Nonlinear systems and nonlinear eatimation theory, in The Mathematics of Filtering and Identifications, M. Hazewin kel and J. C. Williems, eds., Reidel, Dordrecht, The Netherlands, 1981.

(2) R. W. Brockett, Nonlinear control theory and differ ential geometry, Proc. Internat. Congr. Mathematicians, pp. 1357-1368, 1983.

(3) R. W. Brockett and J. M. C. Clark, The geome try of the conditional density functions, in Analysis and Optimization of Stochastic Systems, O. L. R. Jacobs, et. al. eds., Academic Press, New York, 1979.

(4) J. Chen and S. S. -T. Yau, Finite dimensional Filters with nonlinear drift VI: Linear structure of $\Omega$. (preprint)

(5) ——, Finite dimensional filters with nonlinear drift VII: Mitter conjecture and structure of $\eta$, SIAM J. Control and Optimization. (to appear)

(6) J. Chen, S. S. -T. Yau and C. W. Leung, Finite-dimentional filters with nonlinear drift IV: classification of finite-dimensional estimation algebras of maximal rank with state-space dimension 3, SIAM J. Control and Optimization Vol 34, No 1, pp. 179-198, January 1996.

(7) J. Chen, S. S. -T. Yau, and C. W. Leung Finite dimensional Filters with

nonlinear drift VIII, Classfication of finite dimensional estimation algebras of maximal rank with state space dimension 4, (preprint), 1993.

(8) W. L. Chiou, A note on estimation algebras on nonlinear filtering theory, Systems and Control Letters, Vol. 28, pp. 55-63, 1996.

(9) W. L. Chiou and S. S. -T. Yau, Finite-dimensional filters with nonlinear drift II: Brockett's problem on classification of finite-dimensional estimation algebras, SIAM J. Control and Optimization, Vol. 32, No. 1, pp. 297-310, January 1994.

(10) M. H. A. Davis: On a Multiplicative functional trans formation arising in nonlinear filtering theory, Z. Wahrsch. Verw. Gebiete, 54, pp. 125-139, 1980.

(11) R. T. Dong , L. F. Tam. , W. S. Wong. and S. S. -T. Yau, Structure and classification theorems of finite dimensional exact estimation algebras, SIAM J. Control and Optimization, Vol. 29, No. 4, pp. 866-877, July 1991.

(12) S. K. Mitter, On the analogy berween mathematical problrms of nonlinear filtering and quantum physics, Richerche di Automatica, 10 (2): pp. 163-216, 1979.

(13) D. L. Ocone, Finite dimentional estimation algebras in nonlinear filtering, in The Mathematics of Filtring and Identification and Applications. M. Hazewinkel and J. S. Willems. eds. Reidel, Dordrecht, 1981.

(14) L. F. Tam, W. S. Wong and S. S. -T. Yau, On a necessary and sufficient condition for finite dimensionality of estimation algebras, SIAM J. Control and Optimization, Vol. 28, No. 1, pp. 173-185, January 1990.

(15) S. S. -T. Yau, Finite dimensional filters with nonlinear drift I: A class of filters including both Kalman-Bucy filters and Benes filters, Journal of Mathematical Systems, Estimation, and Control, Vol 4, No. 2, pp. 181-203, 1994.

(16) M. Zakai, On the optimal filtering of diffusion processes, Z. Wahrsch. Verw. Gibe., 11, pp. 230-243, 1969.

86年10月28日 收稿
86年11月24日 修正
86年12月10日 接受

# 有關具有最大秩估計代數的一些結果

邱 文 齡
輔仁大學數學系

## 摘 要

Brockett 及 Mitter 分別提出利用估計代數構造有限維的非線性濾過。估計代數的概念扮演瞭解有限維非線性過濾的主要方法。在 1983 年 Brockett 提出將有限維估計代數分類。在 Chen, Yau 及 Leung 的論文中，這三位作者介紹了一矩陣方程式。在系統狀態維度是 4 的情況下，應用此矩陣方程式對最大秩有限維的方程式對最大秩有限維的估計代數作出分類。在本篇論文我們得到（1）此矩陣方程式在狀態空間維度小於 4 的狀況下只有簡易解（trivial-solution）的另證。(2) 有關於此矩陣方程式在狀態空間維度爲任意維的條件下只有簡易解的一個充分必要條件。

**關鍵詞：** 非線性濾過，估計代數。

# 引子選擇及反應緩衝液組成對芥藍幼苗之差異顯示反轉錄－聚合酶連鎖反應譜帶式樣的影響

白佳惠　藍清隆
輔仁大學生物系

## 摘　要

應用差異顯示反轉錄－聚合酶連鎖反應（differential display reverse transcription-polymerase chain reaction；DDRT-PCR）系統，以芥藍幼苗時期的總 RNA，利用 anchored oligo-dT primer 煉合於 mRNA 的 polyadenylated tail 進行反轉錄再和另一條具十個鹼基的逢機引子組合進行聚合酶連鎖反應後，於 6% DNA 定序膠體上依不同片段大小呈現出不同的 mRNA 衍生物。研究中發現總 RNA 需先以 DNase I 處理，而 20$\mu$M dNTP、1.5mM $MgCl_2$濃度是進行 PCR 最適當濃度；並可採用 DNA 銀染（silver staining）技術快速呈現譜帶式樣。$T_{12}$MT 和不同 5'-逢機引子進行 DDRT-PCR 均呈現高分子量、無法區分的式樣，$T_{12}$MA、$T_{12}$MG、$T_{12}$MC 和不同 5'-逢機引子進行 DDRT-PCR 則能呈現 50-80 個譜帶。

## 一、緒　　言

基因組 DNA（genomic DNA）的研究固然可提供很多基因訊息，然而若想瞭解生物生長與分化、環境逆境的適應等生化及生理反應機制相關基因表現的表現和調節，則必須由 mRNA 方面著手；但因 mRNA 只佔總 RNA 量約 1-4%，且其中更有一些屬於稀有 mRNA，mRNA 的種類和表現量也受不同發育時期及不同組織的影響，所以如何達成由 mRNA 建立完整 cDNA 庫（cDNA library）是首先要克服的技術難題。

建立 cDNA 庫，一般常用的建構策略有：1）mRNA 由反轉錄酶利用 Oligo-dT 爲引子合成出第一股 cDNA 後，分別可由 Hairpin priming（Maniatis *et al*., 1982）、Oligo-priming（Land *et al*., 1981）及 replacement synthesis（Okayama and Berg, 1982; Gubler and Hoffman, 1983）法得到 ds-DNA，最後則是利用 dC-tail 或加 linker 的方法來選殖。2）利用 6-mer 的逢機引子及反轉錄酶由 mRNA 的內部合成第一股 cDNA，仍以逢機引子來合成第二股 cDNA，再以 DNA 銜接酶（DNA ligase）銜接而得 ds-cDNA，最後仍可以加 dC-tail 或 linker 來達成選殖目的（Koike *et al*., 1987）。近年來 PCR（polymerase chain reaction）技術的蓬勃發展，已先後有學者經由 SI-PCR（sequence-independent PCR）（Froussard, 1993）建立 cDNA 庫，以解決傳統 cDNA 庫選殖過程中極易遺漏稀有 mRNA 的缺點。

Liang 與 Pardee 則發展出一套差異顯示反轉錄－聚合酶連鎖反應（differential display reverse transcription-polymerase chain reaction, DDRT-PCR）的方法，以比較並找出正常細胞與癌細胞在基因表現上的差異（Liang and Pardee, 1992）。此技術用四條具 2-anchored base 的 3'-引子［$T_{12}MA$、$T_{12}MC$、$T_{12}MG$ 和 $T_{12}MT$，其中 $T_{12}$表示 12 個 dTs，M 則爲 dA、dC、及 dG 退化試劑（degenerate mixtures）］固著於 mRNA 的 3' 端而將整個 mRNA 族群經由反轉錄區分成四個次族群，再配合 3'-引子，經由 10-mer 逢機引子所組成的 5'-引子與新反轉錄出的第一股 cDNA 上離 3' 端不同距離處的互補序列雜合進行 PCR 達到放大的目的，且使不同的 mRNA 種類可以其對應 PCR 產物大小，在 6% DNA 定序膠體上區分出；此法可不須先由總 RNA 中純化出 mRNA 即可得到全由 mRNA 而來的 cDNA；而且如此操作可避免純化 mRNA 過程所造成 mRNA 分解或污染、mRNA 的損失，故亦只需少量的總 RNA；加上 PCR 技術的應用及其譜帶式樣有高達 95% 以上的重覆性（Liang and Pardee, 1992; Liang *et al*., 1993），對建立基因庫的貢獻重大。

芥藍（*Brassica oleracea* var. *alboglabra* Bailey）與其他同爲蔓苔屬的許多作物均爲國內重要蔬菜作物，本研究應用 DDRT-PCR 技術於芥藍幼苗期的 RNA，以銀染法（silver staining）顯現 DNA 定序膠體上由不同的 mRNA 衍生而來的 PCR 產物；後續研究將選擇瓊脂膠體上適當大小片段建立 cDNA 庫、進行 DNA 定序以編目芥藍 cDNA（cDNA cataloguing）。

## 二、材料及方法

圓葉白花芥藍種子由農友種苗公司所提供。3'-端引子爲 $T_{12}MN$（M 爲 G 或 A 或

C，N 爲 G 或 A 或 C 或 T）由美國 Oligos Etc. 公司所合成；5'-端引子乃根據以下三條件（Mou *et al*.，1994）由加拿大英屬哥倫比亞大學生物科技實驗室合成的二套十個鹼基逢機引子組中選擇適當引子：GC 含量百分比爲 50%、幾近無自配對性（self-complementarity）、且 5'-端及 3'-端最後二個鹼基至少含一個 G 或 C。

取 2 g 鮮重的植物組織，以改良自 de Vries *et al*.（1988）法抽取總 RNA。在 20μl 反應體積中，取 2μg 經 DNase I 處理過的芥藍總 RNA，2.5μM 3'-端引子（$dT_{12}$MA、$dT_{12}$MT、$dT_{12}$MC、$dT_{12}$MG 四者其中之一，其中 M 可爲 dA、dG 或 dC），加處理過的 DEPC（Diethylpyrocarbonate）-$H_2O$ 至 10μl，於 70℃ 反應 10min 以使 RNA 變性打開而利於引子的煉合，再加入 4μl 的 5×First-strand buffer（250mM Tris-HCl pH 8.3，375mM KCl，15mM $MgCl_2$）、10mM DTT（Dithiothreitol）、20μM dNTP、20U RNasin ribonuclease inhibitor，於 37℃ 下作用 2min 後，很快的加入 300U M-MLV（Moloney murine leukemia virus reverse transcriptase），於 37℃ 下作用 1hr 後，以 95℃ 作用 5min 後置於冰上待用或儲於 −20℃ 冷凍櫃中。接著取 2.5μM 與反轉錄反應中同一種 3'-端引子、0.5μM 5'-端引子、20μM dNTP、250 μg/ml BSA（bovine serum albumin）、1μl 10×Super Taq buffer（500mM KCl，100mM Tris-HCl pH 9.0，15mM $MgCl_2$，1% Triton X-100）、0.5U Super Taq DNA polymerase、2μl 由反轉錄反應中所合成的第一股 cDNA，補水至 10μl 混合均勻，置於 Air Thermo-Cycler 1605（Idaho Technology）機器進行以下反應。首先是 94℃ 2min，接著是在下列條件下循環 30 次：94℃ 2sec 以使 DNA 變性，42℃ 2sec 使引子煉合（primer annealing），72℃ 30sec 使引子延伸（primer extension）。

利用 Gibco S2 定序電泳系統製備 6% 聚丙烯醯胺膠體（polyacrylamide gel），於上下槽置入 0.6×TBE buffer（75mM Tris base，25mM borate，1.5mM EDTA），以 1600V 預電泳 30min，再以 1900V 60W 進行電泳；將此含膠體之短玻璃放入置有 2 升固定/終止液（10% glacial acetate）的塑膠盆中輕搖 20min，倒走固定/終止液，留待終止顯影反應時使用；以 $ddH_2O$ 三次潤溼膠體，每次 2min；換置入 2 升染色液（12mM $AgNO_3$，0.056% HCOH）輕搖 30min，倒掉 染色液，以 $ddH_2O$ 潤濕膠體 20sec，再換置入 2 升 10-12℃ 預冷的展現液（0.283M $Na_2CO_3$，0.056% HCOH，4μM $Na_2S_2O_3 \cdot 5H_2O$）至譜帶清楚出現爲止。接著加入 2 升固定/終止液，輕搖 2-3min，再以 $ddH_2O$ 兩次潤濕膠體，每次 2 min（Bassam *et al*.，1991），最後風乾即可。

## 三、結　　果

本研究進行 PCR 時使用 Super Taq DNA polymerase (HT Biotechnology)；因大多數已發表文章多爲 40℃ 或 42℃ 並無太大改變，而在本研究之初也已比較過二種煉合溫度，發現其譜帶式樣極相似，故選擇 42℃爲 PCR 的煉合溫度。此外，在 3'-引子選擇方面，選用 $T_{12}MA$、$T_{12}MT$、$T_{12}MG$ 和 $T_{12}MC$；而 5' 引子則選取 50% GC 含量、幾近無自配對性以外、且 5'-端及 3'-端最後二個鹼基至少含一個 G 或 C。圖 1A 顯示 DNA 污染對 DDRT-PCR 的影響，右圖乃總 RNA 不經 DNase I 處理，利用逢機引子 No.1、No.2 和 No.3 (圖示以 1，2，3 區分) 分別和 $T_{12}MA$ 組合進行 DDRT-PCR，由圖可明顯看出疑似有 DNA 污染的總 RNA 即使不經反轉錄反應，僅進行 PCR 也有很多的譜帶產生 (右圖之 1D，2D，3D)，表示這些譜帶極可能是由污染 DNA 進行 PCR 所產生。而圖左中的總 RNA 先以 DNase I 處理，再利用相同的引子組合進行 DDRT-PCR (左圖之 1R，2R，3R)，其譜帶式樣確實已不如右圖之 1R、2R、3R 結果複雜，且若不經反轉錄反應，直接進行 PCR 則不會有任何譜帶產生 (左圖之 1D，2D，3D)，顯示 DNase I 處理確實可解決總 RNA 中的 DNA 污染對 DDRT-PCR 表現式樣所造成的僞陽性 (false positive)。

利用逢機引子 No.1、No.2 及 No.3 分別和 $T_{12}MA$ 組合，於 1.25mM、1.5mM (最適合 Super Taq DNA Polymerase 進行 PCR)、3.0mM 不同鎂離子濃度下進行 DDRT-PCR (圖 1B)，結果顯示不論何種引子組合，均在 1.5mM 濃度有最佳的顯示式樣。本研究最初使用 2$\mu$M dNTP 以進行 PCR，但得到很少譜帶，故繼以逢機引子 No.1、No.2 和 No.3 分別和 $T_{12}MA$ 組合，並於不同 dNTP 濃度 (2$\mu$M、10$\mu$M、20$\mu$M、30$\mu$M) 進行 DDRT-PCR (圖 1C)，如圖所示，不論何種引子組合，在 20$\mu$M 和 30$\mu$M dNTP 濃度條件下可得較好的顯示式樣，2$\mu$M dNTP 濃度下則最差。

一般採用 3' 引子：5' 引子 = 5：1 的濃度比例，主要目的乃使 3'-引子能煉合得較好，避免 5'-引子佔優勢而複製出兩端皆具 5'-引子序列的 DDRT-PCR 產物。圖 2 之右圖，分別由 $T_{12}MA$、$T_{12}MC$、$T_{12}MG$ 與 No.1、No.2、No.3、No.4、No.5 的逢機引子組合，在 10：1 的濃度比例下進行 PCR，結果顯示 $T_{12}MA$ 與不同的逢機引子在 10：1 的濃度比例下確實呈現出不同的譜帶式樣，但 $T_{12}MG$ 及 $T_{12}MC$ 與不同的逢機引子組合，於 10：1 濃度比例下卻產生極相似的譜帶式樣。利用與右圖相同的引子組合，以正常的引子濃度比例 (即 2.5$\mu$M：0.5$\mu$M) 進行 PCR (圖 2 之

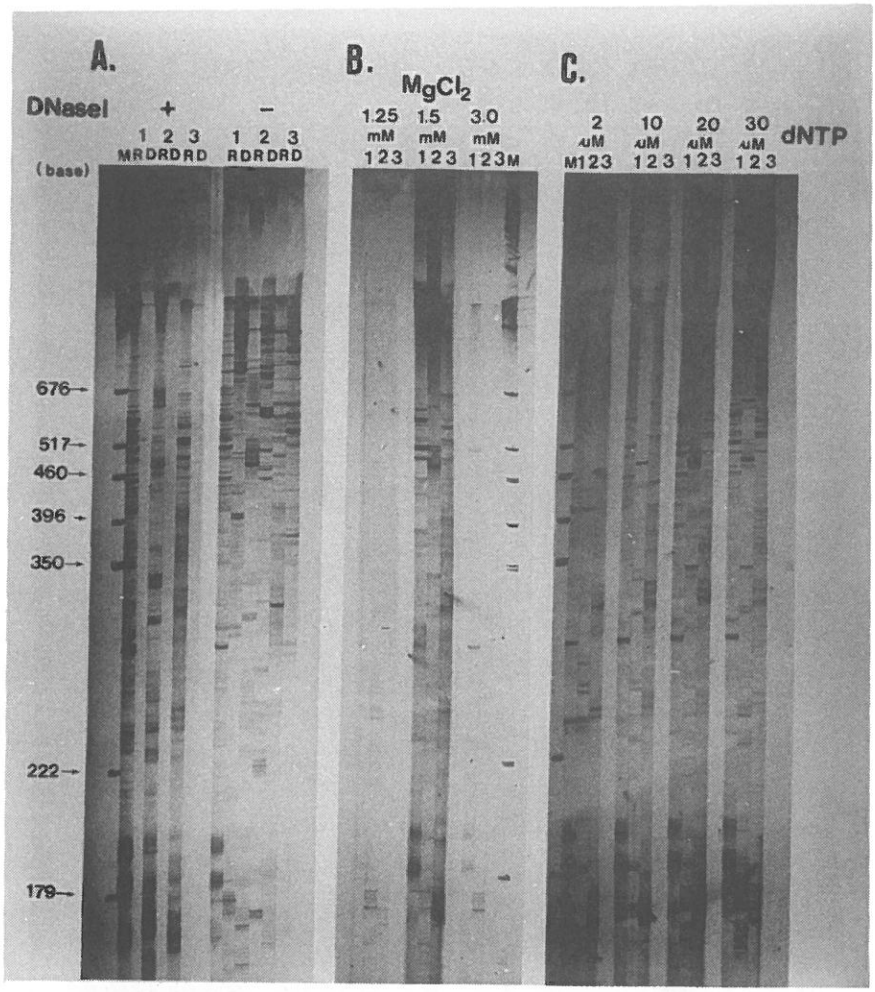



**圖 1** 差異顯示反轉錄－聚合酶連鎖反應（DDTR-PRC）的最適當條件。利用逢機引子 No.1，No.2，No.3 分別與 $T_{12}MA$ 組合（圖上分別以 1，2，3 表示）進行 DDRT-PCR 後，於 6%DNA 定序膠體經銀染呈現。M 爲 pGEM 標記。A. DNase I 處理總 RNA 對表現式樣的影響，R 爲經反轉錄，D 爲不經反轉錄，證實總 RNA 經 DNase I 的處理是必要的。B·鎂離子濃度對表現式樣的影響，PCR 中，鎂離子濃度分別爲 1.25mM，1.5mM，及 3.0mM，結果顯示最適當的濃度爲 1.5mM。C. dNTP 濃度對表現式樣的影響，PCR 中 dNTP 濃度分別爲 2μM，10μM，20μM 及 30μM，結果顯示 20μM dNTP 可呈現最佳表現式樣。

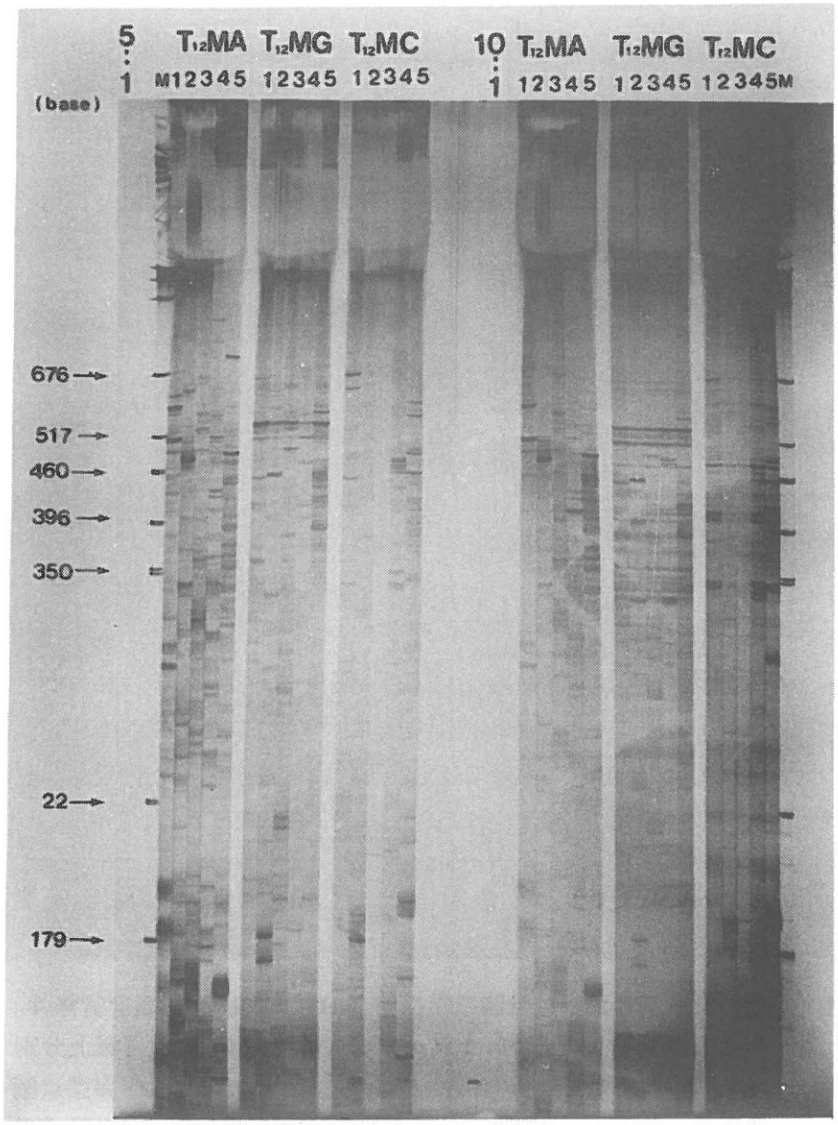



**圖 2　不同 $T_{12}$MN 引子/逢機引子濃度比例對 DDRT-PCR 表現式樣的影響。於 $T_{12}$MN 引子/逢機引子濃度比例爲（左）5：1 和（右）10：1 條件下，分別以 $T_{12}$MA，$T_{12}$MG，$T_{12}$MC 與逢機引子 No.1，No.2，No.3，No.4，No.5（圖上分別以 1，2，3，4，5 表示）組合進行 DDRT-PCR，並於 6%DNA 定序膠體呈現。**

左圖)，結果顯示不論 $T_{12}MA$、$T_{12}MG$ 或 $T_{12}MC$，其與不同逢機引子組合，確實得到不同的譜帶式樣；同時，比較 3' 引子：5' 引子 = 5：1 (左圖) 與 10：1 (右圖) 濃度比例下，發現相同的 $T_{12}MA$ 和逢機引子組合所產生的譜帶式樣是極相似的。

根據本研究所改良的最適當條件，利用 3'-引子 $T_{12}MA$、$T_{12}MT$、$T_{12}MC$ 和 $T_{12}MG$ 分別與 5'-端十條具不同序列的逢機引子組合進行 DDRT-PCR 後，於 6% DNA 定序膠體分離，再經銀染顯色 (圖 3)，除了 $T_{12}MT$ 和不同逢機引子組合進行 DDRT-PCR 產生無法區分的高分子量產物，及 $T_{12}MC$ 和 No.3 逢機引子組合進行 DDRT-PCR 無任何譜帶產生外，其餘引子組合結果平均可呈現出 50-80 條譜帶 (最少者也有約 30 條譜帶左右)；且圖中每一引子組合皆有一控制組 (圖中顯示於每一反應組右側)，結果顯示均無任何污染 DNA 被當成模板複製出來。繼而選殖由 $T_{12}MA$、$T_{12}MC$ 和 $T_{12}MG$ 分別與不同逢機引子組合所得的 DDRT-PCR 產物中 179bp-676bp cDNA 片段入 pUC18 載體，最後以 PCR 快速檢偵測轉殖菌落，並初步挑選含有不同且適當大小外緣 DNA 片段的殖株培養，以供定序分析本研究 DDRT-PCR 技術的可靠性。

## 四、討　　論

影響 DDRT-PCR 顯示式樣的因素包括 DNA 污染、引子選擇、3'-引子/5'-引子濃度比例、PCR 時所用的 DNA 聚合酶種類 (Haag and Roman, 1994)、煉合溫度、鎂離子濃度、dNTP 濃度。本研究在 3'-引子的選擇是根據 Liang 等人於 1993 年提出的修正意見，因 anchored oligo-dT 引子 $T_{12}MN$ 的專一性 是由引子的最後一個鹼基 N 所提供，倒數第二個鹼基有退化 (degeneracy) 現象，故以 $T_{12}MA$、$T_{12}MT$、$T_{12}MG$ 和 $T_{12}MC$ 4 條取代原本的 12 條 $T_{12}MN$ 引子 (Liang *et al.*, 1993)。但由圖 3 所示，$T_{12}MT$ 和不同逢機引子組合進行 DDRT-PCR 產物均表現出模糊式樣，這可能是因 $T_{12}MT$ 爲退化引子組 (degenerate primer)，且主要由引子的最後一個鹼基提供專一性，而此時最後一個鹼基卻爲 dT，故較可能於反轉錄反應時煉合於 mRNA poly A 端的任意處，造成同一種 mRNA 反轉錄複製出各種不同大小的 cDNA；因此只有 $T_{12}MA$、$T_{12}MG$、$T_{12}MC$ 與逢機引子的組合表現預期式樣。Liang 等人又曾將 3'-引子改良成 5'-含 *Hind* III 限制酶切割位的 H-$T_{11}A$、H-$T_{11}G$、H-$T_{11}C$ 以避免當 $T_{12}MN$ 的 M 爲 A 鹼基會有不利煉合的情形發生，並減少 $T_{12}MT$ 成爲 oligo-dT 污染而產生模糊式樣 (Liang *et al.*, 1994)。而在 5'-端逢機引子的選擇是根據 Liang 及

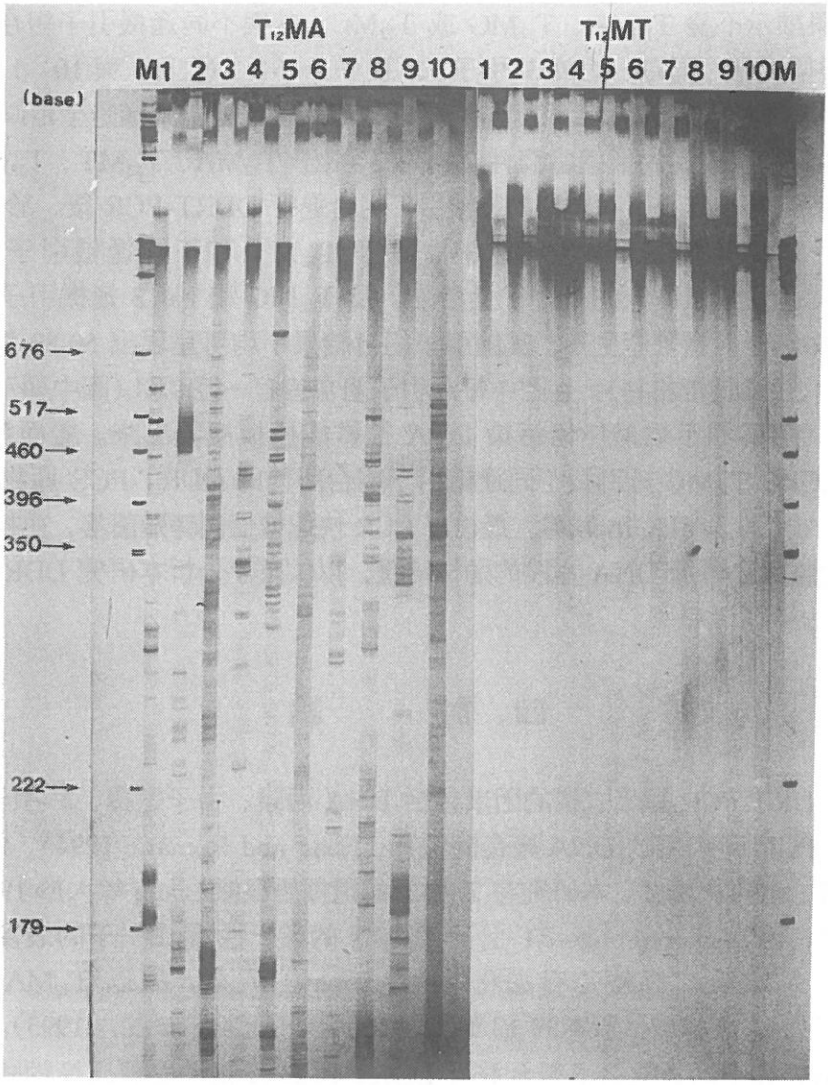
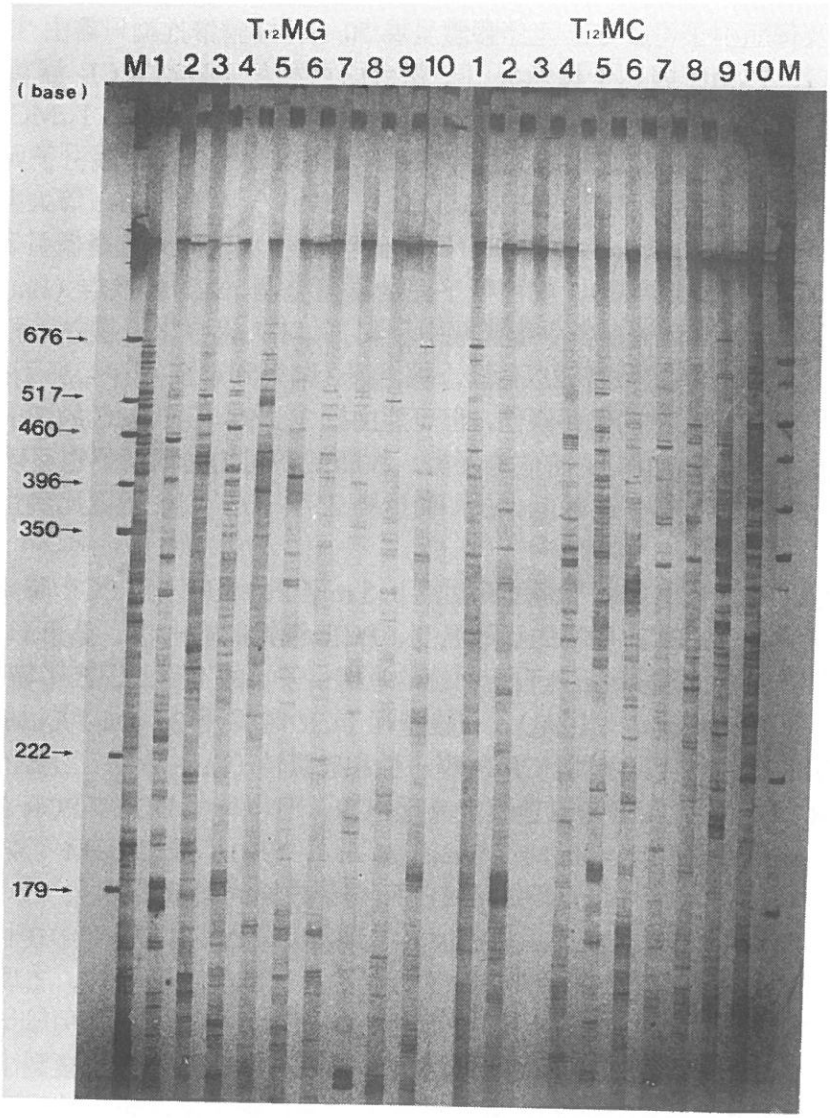



**圖 3　$T_{12}MA$，$T_{12}MT$，$T_{12}MG$ 和 $T_{12}MC$ 分別與十條不同的逢機引子（分別以 1，2，3，4，5，6，7，8，9，10 表示）組合進行 DDRT-PCR 後，於 6% DNA 定序膠體經銀染呈現其表現式樣。圖中每一組反應皆有不經反轉錄的控制組相對於每一反應組的右側，M 爲 pGEM DNA 標記。**




Pardee 最初發表文章中利用 6-mer、7-mer、8-mer、9-mer 及 10-mer 逢機引子與 $T_{12}$MN 組合測試 DDRT-PCR 的結果所顯示，其理論值雖分別可得到 150、38、10、2、小於 1 的譜帶數，但實際上卻反而分別得到 0、0、0、20-30、50-100 的譜帶數，故採用 10-mer 逢機引子以容忍逢機引子與標的序列可有 3-4 個鹼基的錯失配對 (Liang *et al*., 1992; Bauer *et al*., 1993)。雖然 Liang 等人曾根據哺乳動物細胞約有 15000

種 mRNA 及每組引子可於 6% 定序膠體呈現 50-100 條譜帶推論只需由 $T_{12}MA$、$T_{12}MT$、$T_{12}MG$、$T_{12}MC$ 與二十種逢機引子組合即可利用 DDRT-PCR 涵蓋幾乎整個 mRNA 族群。但 Crawford 等人則認爲 $T_{12}MA$、$T_{12}MT$、$T_{12}MG$、$T_{12}MC$ 與二十五種逢機引子組合僅能涵蓋 50% 的 mRNA 族群，而以七十五種逢機引子與其組合也只能涵蓋 90% 的 mRNA 族群 (Crawford *et al.*, 1993)。另外 Bauer 等人則認爲利用十二條 $T_{12}MN$ (M 爲 dA、dG 或 dC，N 爲任意鹼基) 與二十六種逢機引子組合 (每一組合平均呈現 120 條譜帶)，就可幾乎完全涵蓋整個 mRNA 族群 (Bauer *et al.*, 1993)。於本研究中究竟需要多少種逢機引子與 $T_{12}MN$ 組合才能建立一相當完整的 cDNA 庫，因數據未完整致實難以估計。綜合圖 2 結果判定 3' 引子：5' 引子 = 5：1 濃度比例是 DDRT-PCR 的必要條件，但可能由於 $T_{12}MG$ 和 $T_{12}MC$ 的 Tm 較高，所以在 3' 引子：5' 引子 = 10：1 時佔盡優勢，故即使不同的逢機引子也呈現出相似的譜帶式樣，而 $T_{12}MA$ 的 Tm 較低，故即使提高 3' 引子/5' 引子濃度比例也不會對譜帶式樣產生明顯影響。

一般進行 PCR 時，鎂離子濃度範圍爲 0.5 mM-2.5 mM，因 PCR 時鎂離子濃度與引子煉合、模板和 PCR 產物的解離溫度、PCR 產物的專一性、是否形成 primer-dimer (於本研究中已選擇幾近無自配對性的逢機引子，故不受此因素影響)、酶活性及忠實性有關 (Innis *et al.*, 1990)。一般進行 PCR 時採用 20 $\mu$M-200$\mu$M dNTP 濃度，較低的 dNTP 濃度可增加 PCR 的專一性和忠實性 (Innis *et al.*, 1990)；而一般 DDRT-PCR 研究使用 2$\mu$M dNTP 濃度進行 PCR，但也有些 DDRT-PCR 採用 20$\mu$M (Lohmann *et al.*, 1995)、50$\mu$M (Welsh *et al.*, 1990) 或 200$\mu$M (Sokolvo and Prockop, 1994)，本研究中則發現 20$\mu$M 下可得到較好的顯示式樣。

一般的 DDRT-PCR 法均使用放射性 α-[$^{35}$S] dATP，在本研究中提供 DNA 顯示式樣的非放射線呈現方法。銀染法除了可避免產生放射性廢料外，主要的優點是極省時 (只需 1hr 即可顯色)，而顯色後所呈現的譜帶式樣清晰，若用於比較兩種以上不同狀況顯示式樣，更是利於由定序膠體上回收有差異的譜帶並複製出某一特定譜帶，不易受到其他相鄰譜帶的影響。至於由定序膠體上回收銀染特定譜帶，並複製的方法更是簡易且專一性高 (Sanguinetti *et al.*, 1994)，且在發表的文章中也確實以銀染方法做 DDRT-PCR 並回收、複製 (Lohmann *et al.*, 1995)；至於銀染法的缺點則是其靈敏度不如放射性檢測系統，不利於區分由稀少套數的基因所展現的差異。

## 五、誌　　謝

本研究承行政院農業委員會及輔仁大學聖言會單位補助部份研究經費，謹此誌

謝。

## 參考文獻

(1) Bassam, B. J., C. A. Gustavo, and M. G. Peter. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal. Biochem. 196: 80-83.

(2) Bauer, D., H. Muller, J. Reich, H. Riedel, V. Ahrenkiel, P. Warthoe, and M. Strauss. 1993. Identification of differential expressed mRNA species by an improved display technique (DDRT-PCR). Nucleic Acids Res. 21: 4272-4280.

(3) Crawford, D. R., C. A. Edbauer-Nechamen, C. V. Lowry, S. L. Salmon, Y. K. Kim, J. M. S. Davies, and K. J. A. Davies. 1993. Assessing gene expression during oxidative stress. Meth. Enzymol. 234: 175-217.

(4) Froussard, P. 1993. rPCR : A powerful tool for random amplification of whole RNA sequences. PCR Methods Applications 2: 185-190.

(5) Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. Gene 25: 263-269.

(6) Haag, E., and V. Raman. 1994. Effects of primer choice and source of Taq DNA polymerase on the banding patterns of differential display RT-PCR. BioTechniques 17: 226-228.

(7) Innis, M. A., and D. H. Gelfand. 1990. Optimization of PCRs pp. 3-12. In: M. A. Innis, D. H. Gelfand, J. J. Sininsky, and T. J. White. (eds.). PCR Protocols: A Guide to Methods and Applifications. Academic Press, San Diego.

(8) Koike, S., M. Sakai, and M. Muramatsu. 1987. Molecular cloning and characterization of rat estrogen receptor cDNA. Nucleic Acids Res. 15: 2499-2499.

(9) Land, H., M. Grez, H. Hauser, W. Lindenmaier, and G. Schultz. 1981. 5'-Terminal sequences of eukaryotic mRNA can be cloned with high efficiency. Nucleic Acids Res. 9: 2251-2266.

(10) Liang, P., and A. B. Pardee. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971.

(11) Liang, P., L. Averboukh, and A. B. Pardee. 1993. Distribution and cloning

of eukaryotic mRNA by means of differential display: refinements and optimization. Nucleic Acids Res. 21: 3269-3275.

(12) Liang, P., W. Zhu, X. Zhang, Z. Guo, R. P. O'Connell, L. Averboukh, F. Wang, and A. B. Pardee. 1994. Differential display using one-base anchored oligo-dT primers. Nucleic Acids Res. 22: 5763-5764.

(13) Lohmann, J., H. Schickle, and T. C. G. Bosch. 1995. REN display, a rapid and efficient method for nonradioactive differential display and mRNA isolation. BioTechniques 18: 200-202.

(14) Maniatis, R., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. Cold Spring Harbor Laboratory Press, New York.

(15) Mou, L., H. Miller, J. Li, E. Wang, and L. Chalifour. 1994. Improvements to the differential display method for gene analysis. Biochem. Biophys. Res. Commun. 199: 564-569.

(16) Okayama, H., and P. Berg. 1982. High-efficiency cloning of full-length cDNA. Mol. Cell Biol. 2: 161-170.

(17) Sanguinetti, C. J., E. D. Neto, and A. J. G. Simpson. 1994. Rapid silver staining and recovery of PCR products seperated on polyacrylamide gels. BioTechniques 17: 914-922.

(18) Sokolov, B. P., and D. J. Prockop. 1994. A rapid and simple PCR-based method for isolation of cDNAs from differentially expressed genes. Nucleic Acids Res. 22: 4009-4015.

(19) de Vries, S., H. Hoge, and T. Bisseling. 1988. Isolation of total and polysomal RNA from plant tissues. pp. A7: 1-13. In: S. B. Gelvin, R. A. Schilperoort, and D. P. S. Verma. (eds.). Plant Molecular Biology Manual. Kluwer Acad. Publ., Dordrecht.

(20) Welsh, J., J. P. Liu, and A. Efstratiadis. 1990. Cloning of PCR-amplified total cDNA : Construction of a mouse oocyte cDNA library. Genet. Anal. Techn. Appl. 7: 5-17.

86年11月8日 收稿
86年12月15日 修正
86年12月28日 接受

# Effect of primer choice and reaction buffer component on the banding patterns of differentially display RT-PCR (reverse transcriptase-polymerase chain reaction) in Chinese kale seedlings

CHIA-HUI PAI AND CHING-LONG LAN
*Department of Biology*
*Fu-Jen University*
*Taipei, Taiwan 242, R.O.C.*

## ABSTRACT

A DDRT-PCR (differential display reverse transcription-polymerase chain reaction) protocol, starting with total RNA from young Chinese kale seedlings, used an anchored oligo-dT primer to anneal to mRNA at the poly-(A) tail for reverse transcription, and a subsequent PCR with the same anchored oligo-dT primer and a random 10-mer primer, was developed to differentially display Chinese kale mRNA species on a 6% DNA sequencing gel by size. DNA silver staining was used to quickly display the DDRT-PCR products. For the PCR, 20 $\mu$M dNTP and 1.5mM $MgCl_2$ were optimal. In addition, pre-treatment of total RNA with DNase I was necessary.

A gel profile of 50-80 bands was normally displayed when, together with a random 10-mer 5'-primer, with one of $T_{12}MA$, $T_{12}MG$, and $T_{12}MC$ 3'-primers was used, but not with the $T_{12}MT$ series.

# $RbTiOAsO_4$ 單晶的聲聲子研究

薛仲貴　　杜繼舜

輔仁大學物理研究所

## 摘　　要

本報告，是利用布里元光學散射實驗，進行研究 $RbTiOAsO_4$ (RTA)單晶的 LA[001]聲聲子光譜與其溫度的關係。發現當溫度升高時，RTA 的聲聲子頻率顯示出明顯的減弱，直到其鐵電相變溫度 ($T_c$～800 ℃)以上時才趨於定值，這個現象被歸因於聲聲子的"軟模 (soft-mode)"出現。其聲子半寬頻率則是在溫度接近相變溫度時，有明顯的最大值，我們認爲 order-parameter (polarization)的 fluctuations 是這個行爲的主因。藉著使用第拜非同調近似式 (Debye anharmonic approximation )，非耦合的聲子頻率 $\omega_a(0)=44.28$ GHz，第拜溫度 $\Theta=300$ K 及非同調係數 $A=6\times10^{-5}K^{-1}$也被得到。

**關鍵詞：** 布里元光學散射；聲聲子；鐵電相變；第拜非同調近似式。

## 一、簡　　介

$RbTiOAsO_4$ 是屬於非線性光學晶體中，具有分子式爲 $M^{1+}TiOX^{5+}O4$ (M = K, Rb, Tl, Cs and X = P, As ) 晶體家族中的一員[1-6]。這類型的晶體在波長 1.06 及 1.32$\mu$m 釹雷射的測試下，發現其在非線性光學的使用上，擁有較高的危險門檻及較寬的接受角，極適合使用於雷射倍頻 (SHG)及光學參數振盪器 (OPO)。另外，RTA 在波導的應用上也有廣泛的發展。在同一個家族的晶體中，$KTiOPO_4$ ( KTP )是其中較著名，已經成功應用於各個領域的晶體[1-6]。但是，由於 KTP 中的正磷酸根會吸

收 4.3 到 3.5μm 波長範圍的光，以至於對光學振盪器的輸出頻率有很大的限制。相較之下，RTA 對紅外光有較寬波長的透明度（0.35～5.3μm），這些特性讓 RTA 較 KTP 更適於非線性光學的運用。RTA 的鐵電相變溫度大約在 800 ℃。室溫時，RTA 呈鐵電相，晶體爲非中心對稱的點群 $C_{2v}$(mm2) 和空間群 Pna2(Z＝8) 所組成的正交結構。其架構是由各角落連結的 $TiO_6$八面體和 $AsO_4$ 四面體所組成。我們將在這篇論文中，對於下列實驗的結果做出討論，RTA 沿著〔001〕聲子方向的聲聲子頻率（或聲速）、半寬度（或阻尼）和溫度的相依關係。尤其是 RTA 中，聲聲子“軟模”的現象。

## 二、布里元散射的基本原理

布里元散射和拉曼散射都是屬於非彈性散射。布里元散射是由於入射光入射到晶體中，與晶體中的聲聲子作用而產生。其頻率位移範圍約在 0.01～5 $cm^{-1}$之間（拉曼頻率範圍約在 10～1000 $cm^{-1}$之間）。以下簡單介紹其原理。

“聲子”指的就是晶格的振動，其中可分爲聲聲子（acoustic phonon）和光聲子（optical phonon）兩類。光聲子是指晶格做反向運動所產生的振動模，頻率較大。而聲聲子則是指在離子所組成的晶格中，晶格做同向運動所產生的振動模，頻率較小。

聲子是由晶格運動所造成，當晶格振動時，所產生的聲聲子在晶體內的各個晶格之間，以固定的頻率形成駐波，此頻率即是聲聲子的頻率。當光入射到晶體時，就和聲聲子的駐波產生所謂“都卜勒”效應，此時雷射光的頻率會產生頻率位移，這便是布里元散射信號，其中，由於“都卜勒”效應的方向，布里元散射會有小於及大於入射光頻率（頻移相同）的兩個峰，分別稱爲史托克斯散射及反史托克斯散射。

我們在這邊先定義所謂的“自由頻譜寬度”（free spectral range），因爲由“自由頻譜寬度”，可以決定聲子頻率的位移量。我們由干涉方程式中的垂直入射情況可得

$$2d = m\lambda = \frac{mc}{v} \tag{1}$$

這裡 $d$ 是兩面鏡子之間的距離，$v$ 是入射光的頻率，$m$ 表示爲第 $m$ 階的干涉。現在我們考慮，有兩個入射波，頻率各別爲 $v_1$ 及 $v_2$，如果它們在鄰近的 order（$m$ 和 $m+1$）產生建設性干涉時，由（1）式，我們知道

$$2d = mc/v_1 \tag{2}$$

$$2d = (m + 1)c/v_2 \tag{3}$$

由 (3) − (2)，我們可得

$$v_2 - v_1 = \Delta v = \frac{c}{2d}(GH_z) \tag{4}$$

這裡的 $\Delta v$ 就是“自由頻譜寬度 (free spectral range)”[7]。

## 三、實 驗 過 程

$RbTiOAsO_4$ (RTA) 單晶是用 tungstate flux 製程來長成，以 X-ray 繞射來確定晶向，再將之切成沿〔100〕、〔010〕、〔001〕晶面的矩形。晶體大小爲 6.5×5×2

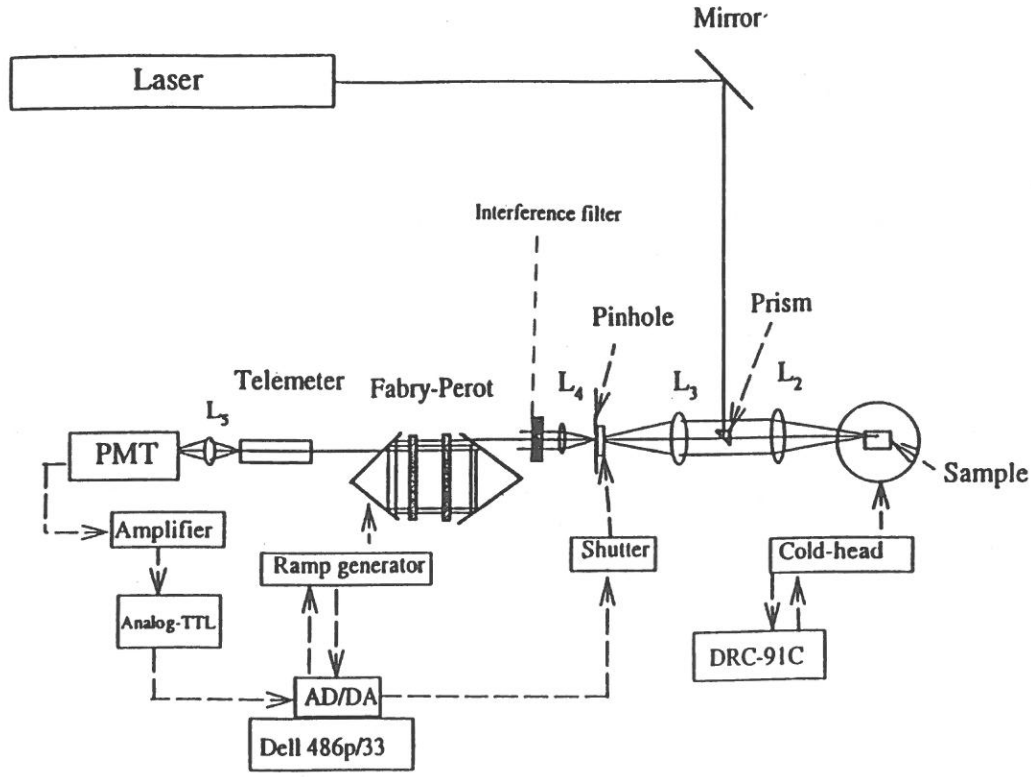



**圖 1　布里元光學 180°散射實驗裝置圖。**

$mm^3$，要測量的表面是利用研磨及抛光使之平滑，以減少來自表面的散射。

實驗裝置如圖 1 所示，其中包括了 coherent 公司的 Innove90-3A 型氬離子雷射，聚焦透鏡組，針孔，濾波片，光閘，Burleigh 公司的 five-pass Fabry- Perot 干涉儀，空間濾波器，光電倍增管（PMT），光子計數系統（Photon Counting Electronic System），高溫爐及控溫器。實驗是採 180°散射的偵測方向，樣品的散射組態爲 Z（XU）Z，"U" 是指我們收集的散射光沒有區別其電場極化（偏振）方向。

我們使用的雷射波長爲 514.5 nm，到達樣品的雷射強度約爲 100 mw，我們將樣品沿（001）聲子方向擺設來得到和波向量（001）縱模聲子（LA）作用的散射光信號。當散射光通過透鏡組及濾光片，濾掉較遠於雷射光頻率的其他散射光（如拉曼散射）之後，進入 Fabry-Perot 干涉儀。利用 Ramp-generator 來微調干涉儀反射鏡的平行。最後聚焦於 PMT 上，藉著計數器將信號轉換，我們可以在電腦上看到實驗的結果。

爲了在實驗中得到更準確的頻率及半寬，我們調整干涉儀鏡子之間的距離 d，來改變自由頻寬，使所得的布里元信號落在第二個雷利散射的範圍內。由石英玻璃的布里元信號，我們決定 FSR 的大小（25.17 GHz）。實驗的數據，是將品置於高溫爐中，將之由室溫加熱到約 850 ℃，溫度的測量是使用 K-type 熱電偶來讀取。

爲了確定布里元信號的位置及半寬，我們將得到的實驗數據，利用下面這個聲子光譜方程式，也就是阻尼簡諧振盪模型的曲線來貼近[8]：

$$S(w) = \frac{\chi_0 \Gamma \omega \omega_0^2}{(\omega^2 - \omega_0^2) + \Gamma^2 \omega^2} \cdot \frac{1}{1 - e^{-\hbar\omega/kT}} \tag{5}$$

$\omega_0$ 和 $\Gamma$ 分別是聲子的頻率和半寬，$\chi_0$ 是晶體的電極化率（susceptibility constant），$k$ 是波茲曼常數，而 $T$ 是絕對溫度。

## 四、結果與討論

由實驗所得到的（001）縱模聲聲子光譜，我們取幾個溫度點中的布里元信號史托克斯線來作圖（如圖 2 所示）。圖 2 中，圓點代表實驗中測到的數據，實線則是利用方程式（5）所貼近的圖形。在貼近的過程中，可以得到各個溫度點，聲子的頻率（$\omega_0$）、半寬（$\Gamma$）、電極化率（$\chi_0$）和背景強度的值。其中，聲子的頻率、半寬與溫度

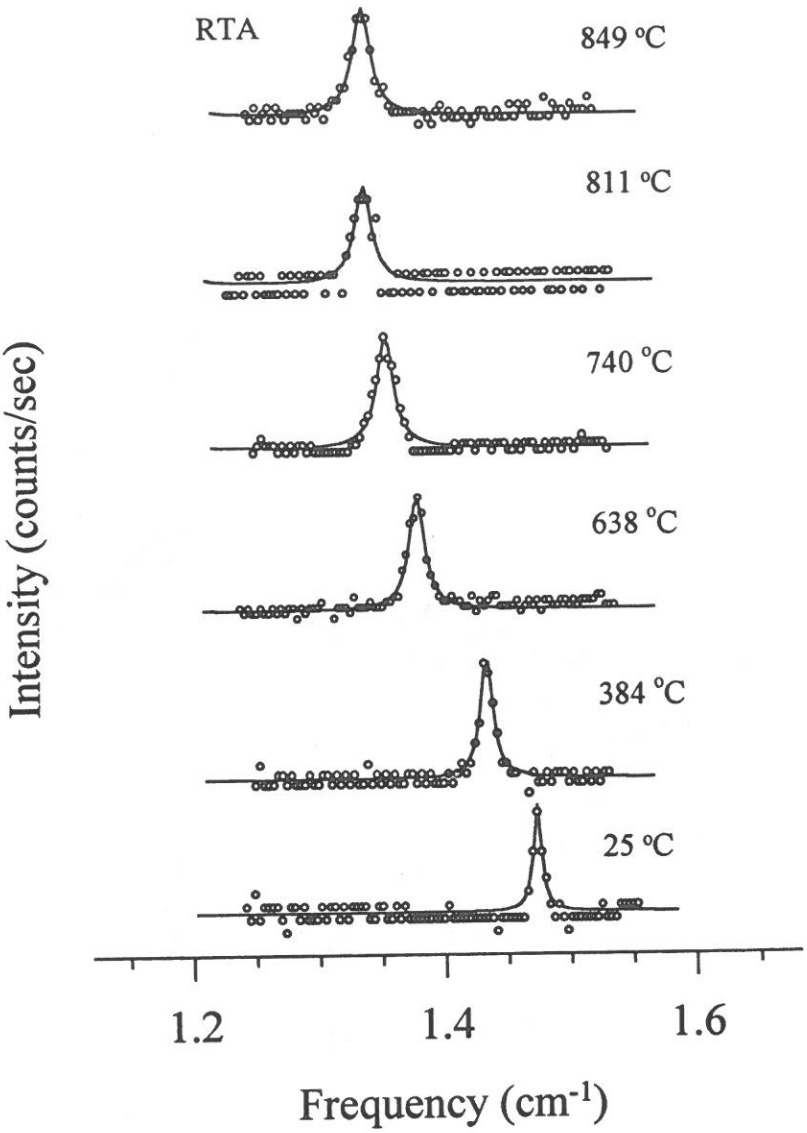



**圖 2　RTA 的 LA〔001〕布里元散射光譜中，聲子的史托克斯散射頻率與溫度的關係圖。圓點是測量的數據，而實線則是藉著（5）式所貼近的曲線。**

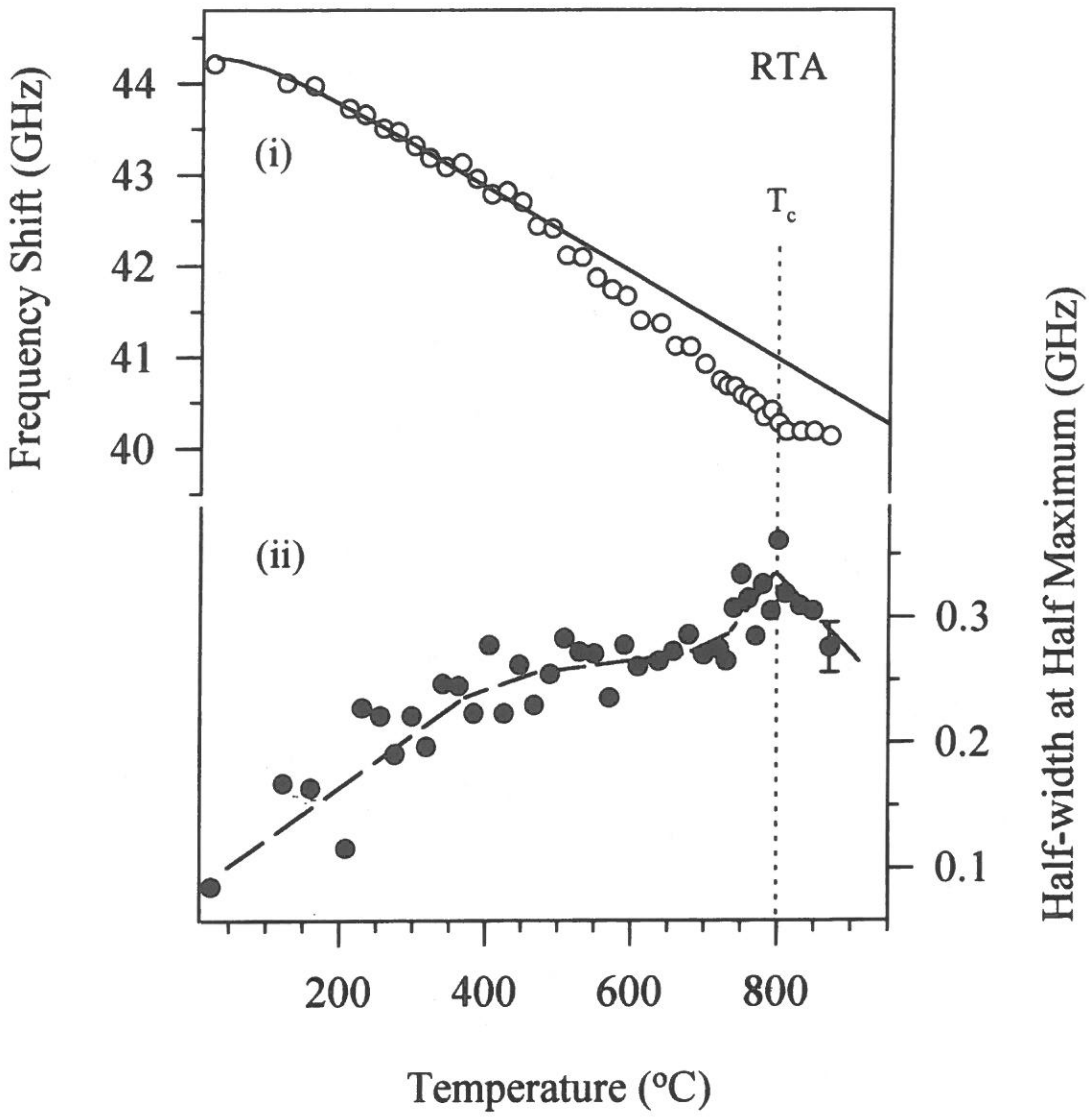



**圖 3　RTA 的 LA〔001〕聲子的頻率（空心圓點）和半寬度的頻率（實心圓點）與溫度的關係圖。**

的關係圖如圖 3 所示。爲了估計 RTA 的耦合效應（coupling），我們藉著貼近（fitting）所測量的較低溫（非耦合效應的區域）的聲子頻率，來計算非耦合聲子的

頻率 $\omega_a$ (T)。非耦合聲子頻率的定義是當聲子遠離相變點時的頻率。非耦合聲子的頻率 $\omega_a$ (T) 與溫度的相依關係可以用以下的第拜非同調近似式 ( Debye anharmonic approximation) 來描述[9]:

$$\omega_a(T) = \omega_a(0)\left[1 - A\Theta F\left(\frac{\Theta}{T}\right)\right] \tag{6}$$

$\Theta$ 是第拜溫度。這裡的 F 是第拜內能方程式[10],

$$F\left(\frac{\Theta}{T}\right) = \frac{3}{(\Theta/T)^4}\int_0^{\Theta/T}\frac{u^3}{e^u - 1}du \tag{7}$$

圖 3 (a) 中的虛線是我們假設溫度是在較靠近室溫，即遠離 RTA 鐵電相變溫度 800 ℃ (非耦合效應的區域) 的情況，將參數 $\Theta$ =300 K, $\omega_a$ (0) =44.28 GHz 和 $A$ =$6\times10^{-5}K^{-1}$帶入 (6) 式中計算，所得的非耦合聲子頻率。從圖 3 (a) 可以看到在大約 500 ℃時，所測量到的頻率漸漸離開所計算的非耦合聲子頻率，也就是聲子軟模的效應開始出現，並在約 800 ℃時最明顯。對一個典型的鐵電材料，常可以看到聲子頻率及半寬的曲線在相變溫度時會有突然的變化 (不連續改變)。所以，我們認爲 RTA 在 800 ℃ 時，發生鐵電相變，也就是由低對稱的鐵電相轉變爲高對稱的順電相。

## 五、結　　論

RTA (001) 縱模聲聲子的頻率與半寬，跟溫度有很大的相關性，在通過鐵電相變點時尤其明顯。在圖 3 中，可以清楚的看到，頻率與半寬都在約 800 ℃時，有明顯轉折現象。由 (6) 式的計算所得的非耦合聲子頻率，可以看到在大約 500 ℃時，所測量到的頻率漸漸離開非耦合聲子頻率，也就是開始有聲子軟模效應的出現。

在這裡，我們要感謝國科會的專題研究計畫補助 NSC-87-2112-M-030-001 及 NSC-86-2112-M-030-002，使這項研究能夠順利完成。

## 參 考 論 文

(1) C. -S. Tu, A. R. Guo, Ruiwu Tao, R. S. Katiyar, Ruyan Guo , and A. S. Bhalla, J. Appl. phys. **79**, 3235 (1996).

(2) A. R. Guo, M. S. thesis, University of Puerto Rico, 1996; A. R. Guo, C. -S. Tu, Ruiwu Tao, R. S. Katiyar, Ruyan Guo, and A . S. Bhalla, Ferroelectric Lett. **21**, 71 (1996).

(3) G. M . Loiacono, D. N. Loiacono and R. A. Stolzenberger, J. Crystal Growth **131**, 323 (1993).

(4) G. M . Loiacono, Appl. phys. lett. **64**, 2457 (1994).

(5) L. T. Cheng, L. K. Chemg, J. D. Bierlein and F. C. Zumsteg, Appl. Phys. Lett. **63**, 2618 (1993).

(6) G. Marnier, B. Boulanger and B Menaert, J. Phys. Cond. Metter**1**, 5509 (1989).

(7) Grant R. Fowles, *Introduction to Modern Optics*, 2nd Ed. (Holt, Rinehart and Winston, INC, New York, 1975), P. 86.

(8) J. F. Ryan, R. S. Katiyar and W. Taylor, Effect Raman Et Th (orie **C2**, 49 (1971).

(9) C. -S. Tu and V. H. Schmidt, Phys. Rev. B **50**, 16167 (1994).

(10) C. Kittel, *Introduction to Solid State Physics*, 5th Ed. (Wiley, New York, 1976), p. 137.

86年11月10日 收稿
86年12月 5日 修正
86年12月21日 接受

# Study of Acoustic Phonons in $RbTiOAsO_4$ Single Crystal

Z. -G. XUE AND C.-S. TU
*Department of Physics,*
*Fu-Jen University,*
*Taipei, Taiwan 242, R.O.C.*

## ABSTRACT

The longitudinal (LA)[001] Brillouin back-scattering phonon spectra have been measured as a function of temperature (25-871 ℃) for $RbTiOAsO_4$ (RTA) single crystal. As temperature increase, the acoustic phonon frequency shows a strong softening (negative coupling) and reaches its a minimum value above the ferroelectric (FE) transition temperature $T_c \sim 800$ ℃. Correspondingly, the half-width of phonon exhibits a broad maximum (near $T_c$) that was attributed to the order-parameter fluctuations. The Debye bare phonon frequency $\omega_a(0)$ Debye temperature $\Theta$, and anhamonicity were also obtained by using the Debye anharmonic approximation.

***Key Words:*** Brillouin scattering; acoustic phonon; ferroelectric transition; Debye anharmonic approximation.

# Asymptotic distributions of the Estimators of the Vännman Type indices for Non-Normal Processes

SY-MIEN CHEN
*Department of Mathematics*
*Fu Jen University,*
*Taipei, Taiwan 242, R.O.C.*

## ABSTRACT

Vännman (1995a) proposed a family of process capability indices $C_p(a,b)$ and natural estimators assuming the knowledge that the measurements of the process follow a normal distribution. Vännman (1995c) derived the asymptotic distribution of the natural estimator under normality. This paper examines some asymptotic properties related to the natural estimator of the Vännman family of indices $C_p(a,b)$ under some regularity conditions. In addition, the assumption of normality is released.

## INTRODUCTION

Process capability analysis has received extensive attention since Burr's (1976) pioneering work in applications of statistical process control to continuously improve quality and productivity. Process capability index, a unitless measurement, is designed to compute the capability of producing conforming items for processes. It is a function of the variance of the process, allowing one to compare the capabilities of different processes even when they are in different measurement scales.

Much research has been performed regarding process capability indices of $C_p$, $C_{pk}$, $C_{pm}$ and $C_{pmk}$. Notable examples are Sullivan (1984), Kane (1986), Chan et al. (1988), Chan et al. (1990), Clements (1989), Chou and Owen (1990), Spiring (1991), Rodriguez (1992), Pearn et al. (1992), Kotz and Johnson (1993), and Chen and Hsu (1995).

For processes involving symmetric tolerances, Vännman (1995a) proposed a superstructure of indices involving two auxiliary parameters, defined for the case of two-sided specification intervals. Such a family of indices generalizes the four basic indices, $C_p, C_{pk}, C_{pm}$ and $C_{pmk}$. She suggested estimators of the proposed indices, and in addition to analytical derivations of the expected value, the mean square error, and the variance of the estimators. Under the assumption of normality, Vännman and Kotz (1995b) provided an explicit form of the distribution of the estimated indices $\hat{C}_p(a,b)$ of the Vännman family of indices $C_p(a,b)$. The above results are not analytically tractable. Also, Vännman and Kotz (1995c) discussed the family's asymptotic expected value and asymptotic mean square error.

In general, if the measurements of a process characteristic follow a non-normal distribution (such a process is called a non-normal process), the validity of any process index or if any index should be calculated remains doubtful. Unfortunately, non-normal processes exist and may frequently go undetected. Some large-sample properties of process capability indices apply to a wide range of process distributions, thereby contributing to our knowledge of behaviour of PCIs under non-normal conditions. Hence, with a clear understanding of their limitations, asymptotic properties can provide valuable insight into the nature of indices.

This paper concentrates primarily on some asymptotic properties of the class of estimators $\hat{C}_p(a,b)$ of the Vännman type indices $C_p(a,b)$ for non-normal processes due to the reasons given above. Concluding remarks are finally made in Section 4.

## NOTATION

$X \xrightarrow{P} Y$ : $X$ converges to $Y$ in probability.

$X \xrightarrow{L} Y$ : $X$ converges to $Y$ in distribution.

$X \xrightarrow{w.p.1} Y$ : $X$ converges to Y with probability 1.

$(W_{1n}, \cdots, W_{kn})$ : A sequence of random vectors.

$(V_{1n}, \cdots, V_{kn})$ : A sequence of random vectors.

$\mu_k = E(X-\mu)^k$ : The $k^{th}$ central moment of the underlying distribution $G$.

$$sgn(a) = \begin{cases} 1, & \text{if } a > 0, \\ 0, & \text{if } a < 0 \end{cases}.$$

$k = O(n^{-2})$, $n \to \infty$ : $|k/n^{-2}|$ remains bounded *as* $n \to \infty$.

## MAIN RESULTS OF THE DERIVATION

Let $X_1, \ldots, X_n$ be a sequence of i.i.d. random variables of measurement from a process which has distribution $G$ with mean $\mu$ and positive variance $\sigma^2$ under stationary controlled conditions.

Let $L$ , $U$ be the lower and the upper specification limit of the measurement of the characteristic in which we are interested. Denote $d = \frac{U\text{-}L}{2}$, half the length of the specification interval [L, U]; $M = \frac{U+L}{2}$, the midpoint of the specfication interval; $T$ is the target value.

The Vännman family of process capability indices $C_p(a,b)$, which depend on two non-negative parameters, a and b, is defined as follows (Vännman (1995a)):

$$C_p(a,b) = \frac{d - a|\mu - M|}{3\sqrt{\sigma^2 + b(\mu - T)^2}},$$

where $a, b \geq 0$.

It is easy to see that $C_p(0,0) = C_p$, $C_p(0,1) = C_{pm}$, $C_p(1,0) = C_{pk}$ , and $C_p(1,1) = C_{pmk}$.

When both the mean $\mu$ and the variance $\sigma^2$ of the measurement are unknown, an estimator considered is defined by

$$\hat{C}_p(a,b) = \frac{d - a|\overline{X}_n - M|}{3\sqrt{S_n^2 + b(\overline{X}_n - T)^2}},$$

where $\overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$ , and $S_n^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X}_n)^2}{n}$ , be the sample mean, the sample variance, respectively.

**Theorem 3.1**

$\hat{C}_p(a,b)$ is a consistent estimator of $C_p(a,b)$ .

[Proof]: Since $\hat{C}_p(a,b)$ is a continuous function of $S_n^2$ and $\overline{X}_n$ , by the fact that $S_n^2 \xrightarrow{P} \sigma^2$ , $\overline{X}_n \xrightarrow{P} \mu$ , we have $\hat{C}_p(a,b) \xrightarrow{P} C_p(a,b)$ , therefore $\hat{C}_p(a,b)$ is a

consistent estimator of $C_p(a,b)$.

**Lemma 3.1**

a) If $(W_{1n},\cdots,W_{kn}) \xrightarrow{L} (W_1,\cdots,W_k)$, and $(V_{1n},\cdots,V_{kn}) \xrightarrow{P} (V_1,\cdots,V_k)$, then $(V_{1n}W_{1n},\cdots,V_{kn}W_{kn}) \xrightarrow{L} (V_1W_1,\cdots,V_nW_n)$.

b) If $(W_{1n},\cdots,W_{kn}) \xrightarrow{L} (W_1,\cdots,W_k)$, and $g$ is continuous with probability 1, then $g(W_{1n},\cdots,W_{kn}) \xrightarrow{L} g(W_1,\cdots,W_k)$.

[Proof]: See page 24 of Serfling (1980).

**Lemma 3.2**

If $\mu_4$ exists, then

a) $(\overline{X}_n, S_n^2)$ is asymptotic normal distributed with asymptotic mean $(\mu,\sigma^2)$, and asymptotic variance $\frac{\Sigma}{n}$, where

$$\Sigma = \begin{bmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{bmatrix}.$$

b) $(\overline{X}_n,\overline{X}_n,S_n^2)$ is $AN((\mu,\mu,\sigma^2), \frac{\Sigma^*}{n})$, where

$$\Sigma^* = \begin{bmatrix} \sigma^2 & \sigma^2 & \mu_3 \\ \sigma^2 & \sigma^2 & \mu_3 \\ \mu_3 & \mu_3 & \mu_- \sigma^4 \end{bmatrix},$$

[Proof]: a) See page 72 of Serfling (1980).

b) By Lemma 2.5 of Chen and Hsu (1995).

**Lemma 3.3**

a) Assume that $\mathbf{X}_n = (X_{n1}, \cdots, X_{nk})$ is $AN(\boldsymbol{\mu}, b_n^2 \Sigma)$, with $\Sigma$ a covariance matrix and $b_n \to 0$. Let $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \cdots, g_m(\mathbf{x}))$, $\mathbf{x} = (x_1,\cdots,x_k)$ be a vector-valued function, where each component function $g_i(\mathbf{x})$ is real-valued and has a nonzero differential $g_j(\mu, \mathbf{t})$, $\mathbf{t} = (t_1, \cdots, t_k)$, at $\mathbf{x}=\boldsymbol{\mu}$. Define $\mathbf{D} = \left[\frac{\partial g_i}{\partial x_j}\middle| \mathbf{x}=\boldsymbol{\mu}\right]_{m\times k}$. Then, $\mathbf{g}(\mathbf{X}_n)$ is $AN(\mathbf{g}(\boldsymbol{\mu}), b_n^2 \mathbf{D}\Sigma\mathbf{D}')$.

b) Let $\{\mathbf{u}(n)\}$ be a sequence of m-component random vectors and $\mathbf{b}$ a fixed

vector such that $\sqrt{n}\ [\mathbf{u}(n) - \mathbf{b}]$ has a limiting distribution $\mathbf{N(0, T)}$ as $n \to \infty$. Let $\mathbf{f(u)}$ be a vector-valued function of $\mathbf{u}$ such that each component $f_j(\mathbf{u})$ has a nonzero differential at $\mathbf{u}=\mathbf{b}$, and let $\left.\frac{\partial f_j(\mathbf{u})}{\partial u_i}\right|_{\mathbf{u}=\mathbf{b}}$ be the i, j th component of $\mathbf{\Phi_b}$. Then $\sqrt{n}\ \{\mathbf{f}(\mathbf{u}(n)) - \mathbf{f(b)}\}$ has a limiting distributing distribution $\mathbf{N(0, \Phi' T \Phi)}$.

[Proof]: a) See page 122 of Serfling (1980).

b) See Theorem 4.2.3 in T.W. Anderson (1984).

**Theorem 3.2**

If $\mu_4$ exists, then

$$\sqrt{n}\ (\hat{C}_p(a,b) - C_p(a,b)) \xrightarrow{L} \begin{cases} N(0,\ \sigma_A^2) & , \text{ if } \mu \neq M, \\ -\dfrac{|W_1|}{3\sqrt{\sigma^2 + b\ (\mu - T)^2}} - \dfrac{W_2 d}{6\ [\sigma^2 + b \cdot (\mu - T)^2]^{\frac{3}{2}}}, & \text{if } \mu = M. \end{cases}$$

where

$$\sigma_A^2 = \frac{a^2\sigma^2}{9[\sigma^2 + b(T-\mu)^2]} + sgn(M-\mu)\left[\frac{2ab(T-\mu)\sigma^2 - a\mu_3}{3[\sigma^2 + b(T-\mu)^2]^{\frac{3}{2}}}\right] C_p(a,b)$$
$$+ \left[\frac{b^2(T-\mu)^2\sigma^2 - b(T-\mu)\mu_3 + (\mu_4 - \sigma^4)/4}{[\sigma^2 + b(T-\mu)^2]^2}\right] C_p^2(a,b);$$

and $(W_1, W_2) \sim N((0,0),\ \mathbf{H}\Sigma^*\mathbf{H}')$, where $\mathbf{H} = \begin{bmatrix} a & 0 & 0 \\ bv - 2bT & bu & 1 \end{bmatrix}$.

[Proof]: Define

$$g(u,v) = \left[1 - \frac{a|M-u|}{d}\right] \frac{d}{3\sqrt{v + b(T-u)^2}},$$

for $u \in (L, U)$, $v > 0$, then $C_p(a,b) = g(\mu, \sigma^2)$. And $\sqrt{n}\ (\hat{C}_p(ab) - C_p(a,b))$ $= \sqrt{n}[g(\overline{X}_n, S_n^2) - g(\mu, \sigma^2)]$.

(1)When $L < \mu < M$, $L < u < M$

Since

$$g(u,v) = \left[1 - \frac{a(M-\mu)}{d}\right]\left[\frac{d}{3\sqrt{v + b(T-\mu)^2}}\right],\ L < u < M, v > 0, \text{ then}$$
$$\frac{\partial g}{\partial u}(u,v) = \frac{a}{3\sqrt{v + b(T-u)^2}} + \left[1 - \frac{a(M-u)}{d}\right]\left[\frac{bd(T-u)}{3[v + b(T-u)^2]^{3/2}}\right],$$

and

$$\frac{\partial g}{\partial v}(u,v) = \left[1 - \frac{a(M-u)}{d}\right]\left[-\left(\frac{d}{6[v+b(T-u)^2]^{\frac{3}{2}}}\right)\right].$$

Moreover, if one define

$$D = \left(\frac{\partial g}{\partial g}\Big|_{\mu,\sigma^2}, \frac{\partial g}{\partial v}\Big|_{\mu,\sigma^2}\right),$$

then $D \neq (0,\ 0)$, and by Lemma 3.2a and Lemma 3.3b, we have $\sqrt{n}\ [g(\overline{X}_n, S_n^2) - g(\mu,\sigma^2)] \xrightarrow{L} N(0,\ \sigma_{A1}^2)$, where

$$\sigma_{A1}^2 = D\,\Sigma\,D' = \frac{a^2\sigma^2}{9[\sigma^2+b(T-\mu)^2]} + \left[\frac{2ab(T-\mu)\sigma^2 - a\mu_3}{3[\sigma^2+b(T-\mu)^2]^{\frac{3}{2}}}\right]C_p(a,b)$$
$$+\left[\frac{b^2(T-\mu)^2\sigma^2 - b(T-\mu)\mu_3 + (\mu_4-\sigma^4)/4}{[\sigma^2+b(T-\mu)^2]^2}\right]C_p^2(a,b)$$
$$= \frac{a^2\sigma^2}{9[\sigma^2+b(T-\mu)^2]} + sgn(M-\mu)\left[\frac{2ab(T-\mu)\sigma^2 - a\mu_3}{3[\sigma^2+b(T-\mu)^2]^{\frac{3}{2}}}\right]C_p(a,b)$$
$$+\left[\frac{b^2(T-\mu)^2\sigma^2 - b(T-\mu)\mu_3 + (\mu_4-\sigma^4)/4}{[\sigma^2+b(T-\mu)^2]^2}\right]C_p^2(a,b);$$

(2)When $M < \mu < U$, $M < u < U$

Since

$$g(u,v) = \left[1 + \frac{a(M-u)}{d}\right]\left[\frac{d}{3\sqrt{v+b(T-u)^2}}\right], M < u < U, v > 0,$$

is a real valued function and is differentiable for all $u \in (L,U)$, $v > 0$, then

$$\frac{\partial g}{\partial u}(u,v) = \frac{-a}{3\sqrt{v+b(T-u)^2}} + \frac{[d-a(-M+u)]b(T-u)}{3[v+b(T-u)^2]^{\frac{3}{2}}},$$

and

$$\frac{\partial g}{\partial v}(u,v) = \frac{-[d+a(M-u)]}{6[v+b(T-u)^2]^{\frac{3}{2}}}.$$

Again, define

$$D = \left(\frac{\partial g}{\partial u}\Big|_{\mu,\sigma^2}, \frac{\partial g}{\partial v}\Big|_{\mu,\sigma^2}\right),$$

then $D \neq (0,0)$ , and by Lemma 3.2a and Lemma 3.3b, we have $\sqrt{n}[g(\overline{X}_n, S_n^2) - g(\mu,\sigma^2)] \xrightarrow{L} N(0,\sigma_{A2}^2)$, where

$$\sigma_{A2}^2 = D \sum D' = \frac{a^2\sigma^2}{9[\sigma^2 + b(T-\mu)^2]} - \left[\frac{2ab(T-\mu)\sigma^2 - a\mu_3}{3[\sigma^2 + b(T-\mu)^2]^{\frac{3}{2}}}\right] C_p(a,b)$$
$$+ \left[\frac{b^2(T-\mu)^2\sigma^2 - b(T-\mu)\mu_3 + (\mu_4 - \sigma^4)/4}{[\sigma^2 + b(T-\mu)^2]^2}\right] C_p^2(a,b)$$
$$= \frac{a^2\sigma^2}{9[\sigma^2 + b(T-\mu)^2]} + sgn(M-\mu)\left[\frac{2ab(T-\mu)\sigma^2 - a\mu_3}{3[\sigma^2 + b(T-u)^2]^{\frac{3}{2}}}\right] C_p(a,b)$$
$$+ \left[\frac{b^2(T-\mu)^2\sigma^2 - b(T-\mu)\mu_3 + (\mu_4 - \sigma^4)/4}{[\sigma^2 + b(T-\mu)^2]^2}\right] C_p^2(a,b);$$

(3)If $\mu = M$ ,

$$\sqrt{n}(\hat{C}_p(a,b) - C_p(a,b))$$
$$= \sqrt{n}\left\{\left[1 - \frac{a|\mu - \overline{X}_n|}{d}\right] \frac{d}{3\sqrt{s_n^2 + b(\overline{X}_n - T)^2}} - \frac{d}{3\sqrt{\sigma^2 + b(\mu - T)^2}}\right\}$$
$$= -\frac{a\sqrt{n}|\mu - \overline{X}_n|}{3\sqrt{s_n^2 + b(\overline{X}_n - T)^2}}$$
$$- \frac{\sqrt{n}d\{S_n^2 - \sigma^2 + b[\overline{X}_n^2 - \mu^2 - 2T(\overline{X}_n - \mu)]\}}{3\sqrt{S_n^2 + b(\overline{X}_n - T)^2}\sqrt{\sigma^2 + b(\mu - T)^2}(\sqrt{S_n^2 + b(\overline{X}_n - T)^2} + \sqrt{\sigma^2 + b(\mu - T)^2})}.$$

Let

$$(V_{1n}, V_{2n}) = \left(\frac{1}{3\sqrt{S_n^2 + b(\overline{X}_n - T)^2}},\right.$$
$$\left.\frac{d}{3\sqrt{S_n^2 + b(\overline{X}_n - T)^2}\sqrt{\sigma^2 + b(\mu - T)^2}(\sqrt{S_n^2 + b(\overline{X}_n - T)^2} + \sqrt{\sigma^2 + b(\mu - T)^2})}\right),$$

$$(W_{1n}, W_{2n}) = \sqrt{n}(a(\overline{X}_n - \mu), b[-2T(\overline{X}_n - \mu) + (\overline{X}_n^2 - \mu^2)] + S_n^2 - \sigma^2).$$

Since $S_n \xrightarrow{P} \sigma$ and $\overline{X}_n \xrightarrow{P} \mu$, imply that

$$(V_{1n}, V_{2n}) \xrightarrow{P} \left(\frac{1}{3\sqrt{\sigma^2 + b(\mu - T)^2}}, \frac{d}{6[\sigma^2 + b(\mu - T)^2]^{\frac{3}{2}}}\right).$$

Define $h(u,v,w) = (au, b[-2Tu + uv] + w)$ , which is differentiable. Then

$(W_{1n}, W_{2n}) = \sqrt{n}(h(\overline{X}_n, \overline{X}_n, S_n^2) - h(\mu, \mu, \sigma^2))$. By Lemma 3.2b and 3.3a, $h(\overline{X}_n, \overline{X}_n, S_n^2)$ *is asymptotic normal distributed with asymptotic mean* $h(\mu, \mu, \sigma^2)$, and asymptotic variance $\mathbf{H}\Sigma^*\mathbf{H}'\frac{1}{n}$).

Hence

$(W_{1n}, W_{2n}) \xrightarrow{L} (W_1, W_2)$, where $(W_1, W_2) \sim N((0,0), \mathbf{H}\Sigma^*\mathbf{H}')$, and

$$H = \begin{bmatrix} a & 0 & 0 \\ bv - 2bT & bu & 1 \end{bmatrix}.$$

By Lemma 3.1 (a),

$$(V_{1n}W_{1n}, V_{2n}W_{2n}) \xrightarrow{L} (V_1W_1, V_2W_2).$$

Define $r(u, v) = -|u| - v$, then $r(u, v)$ is a continuous function of $(u, v)$. By Lemma 3.1 (b),

$$r(V_{1n}W_{1n}, V_{2n}W_{2n}) \xrightarrow{L} r(V_1W_1, V_2W_2).$$

Recall

$$r(V_{1n}W_{1n}, V_{2n}W_{2n}) = \sqrt{n}(\hat{C}_p(a, b) - C_p(a, b))$$

and

$$r(V_1W_1, V_2W_2) = \frac{-|W_1|}{3\sqrt{\sigma^2 + b(\mu - T)^2}} - \frac{W_2 d}{6[\sigma^2 + b(\mu - T)^2]^{\frac{3}{2}}}.$$

Therefore,

$$\sqrt{n}(\hat{C}_p(a, b) - C_p(a, b)) \xrightarrow{L} \frac{-|W_1|}{3\sqrt{\sigma^2 + b(\mu - T)^2}} - \frac{W_2 d}{6[\sigma^2 + b(\mu - T)^2]^{\frac{3}{2}}}. \quad \square$$

**Lemma 3.4**

Let $X_1$, $X_2$, $\cdots$ be a sequence of i.i.d. random variables with distribution function $G$. For a positive integer $k$, let $a_k$ be the $k^{th}$ moment and $\mu_k$ be the $k^{th}$ central moment of $G$, let $a_k$ be the sample $k^{th}$ moment and $m_k$ be the sample $k^{th}$ central moment. Then

(1) $a_k \xrightarrow{w.p.1} \alpha_k, m_k \xrightarrow{w.p.1} \mu_k$

(2) $E(m_k) - \mu_k = \dfrac{\frac{1}{2}k(k-1)\mu_{(k-1)}\mu_2 - k\mu_k}{n} + \bigcirc(n^{-2}) n \to \infty$

(3) $M_3 = \dfrac{n^2}{(n-1)(n-2)} m_3$ is an unbiased consistent estimator of $\mu_3$

(4) $M_4 = \dfrac{n(n^2-2n+3)}{(n-1)(n-2)(n-3)} m_4 - \dfrac{3n(2n-3)}{(n-1)(n-2)(n-3)} m_2^2$ is an unbiased consistent estimator of $\mu_4$.

[Proof]: See p69 of Serfling (1980).

**Theorem 3.3**

$\hat{\sigma}_A^2$ is a consistent estimator of $\sigma_A^2$, wher $\hat{\sigma}_A^2 = \dfrac{a^2}{9[1+b\lambda^2]} + sgn(M - \overline{X})$

$$\frac{2ab\lambda - a\dfrac{M_3}{S_n^3}}{3[1+b\lambda^2]^{\frac{3}{2}}}\hat{C}_p(a,b) + \frac{b^2\lambda^2 - b\lambda\dfrac{M_3}{S_n^3} + \left(\dfrac{M_4}{S_n^4} - 1\right)/4}{[1+b\lambda^2]^2}\hat{C}_p^2(a,b),$$

$\lambda = \dfrac{-\overline{X}_n + T}{S_n}$, and $\sigma_A^2$ is defined in Theorem 3.2.

[Proof]: By Theorem 3.1, Lemma 3.4 and Slutsky's theorem [Loève, M. (1978)].

**Theorem 3.4**

$$\frac{\hat{C}_p(a,b) - C_p(a,b)}{\hat{\sigma}_A/\sqrt{n}} \xrightarrow{L} N(0,\ 1)$$

[Proof]: Since $\dfrac{\hat{C}_p(a,b) - C_p(a,b)}{\hat{\sigma}_A/\sqrt{n}} = \dfrac{\hat{C}_p(a,b) - C_p(a,b)}{\hat{\sigma}_A/\sqrt{n}} \dfrac{\sigma_A}{\hat{\sigma}_A}$, by Theorem 3.2 and Slutsky's Theorem, the result follows.

## CONCLUSIONS

Another estimator of the class of indices of $C_p(a,b)$ proposed byVännman is given by

$$\hat{C}_{p,n-1}(a,b) = \frac{d - a|\overline{X}_n - M|}{3\sqrt{S_{n-1}^2 + b(\overline{X}_n - T)^2}},$$

where $S_{n-1}^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}$. Clearly, $\hat{C}_{p,n-1}(a,b) = \sqrt{\frac{n-1}{n}}\hat{C}_p(a, \frac{n-1}{n} \cdot b)$ (Vännman (1995a)). Hence, all the asymptotic properties of $\hat{C}_{p,n-1}(a,b)$ are inherited.

Since Vännman family of indices generalized the four basic indices, $C_p, C_{pk}, C_{pm}$, and $C_{pmk}$, the results in Chan et al. (1990) and Chen and Hsu (1995) are special cases of ours.

Pearn et al. (1992) pointed out some undesirable properties of $C_{pm}$ if the target value $T$ is between $L$ and $U$, but is not equal to $M$, the midpoint of the specification interval. When one has a two-sided specification interval, the case when $T = M$ is quite common for practical situations. Hence, in all Vännman's papers, she emphasized that her discussions were restricted to the case when processes have symmetric tolerances, i. e. $T = M$. Our discussion refutes such an assumption. Furthermore, we do not make any assumption regarding the underlying distribution of the process measurements except that the fourth central moment must exist. This makes the asymptotic results derived in this paper more flexible.

## BIBLIOGRAPHY

(1) Anderson, T. W. (1984). "An Introduction to Multivariate Statistics Analysis," second edition, John Wiley and Sons, New York.

(2) Burr I.R. (1976). "Statistical Quality Control Methods". Marcel Dekker, New York, 234-334.

(3) Chan, L.K., Cheng, S.W. and Spring, F.A. (1988). "New Measure of Process Capability Indices: $C_{pm}$," *Journal of Quality Technology*, 20, 162-175.

(4) Chan, L., Xiong, Z.and Zhang, D. (1990). "On the Asymptotic Distributions of Some Process Capability Indices," *Communication in Statistics-Theory and Methods*, 19 (1), 11-18.

(5) Chen, S.M., and Hsu, N.F. (1995). "The Asymptotic Distributions of the Estimated Process Capability Index $C_{pmk}$," *Communication in Statistics-Theory and Methods*, 24 (5), 1279-1291.

(6) Chou, Y.M. and Owen, D.B. (1990). "A Study of a new Process Capability Indices," *Communication in Statistics-Theory and Methods*, 19 (4), 1231-1245.

(7) Kane, V.E. (1986). "Process Capability Indices," *Journal of Quality Technology*, 18, 41-52.

(8) Kotz, S. and Johnson, N. (1993). "Process Capability Indices," *Chapman & Hall*.

(9) Loève, M. (1978). *Probability Theory* Ⅱ, 4th edition, Springer-Verlag, Berlin and New York.

(10) Pearn, W., Kotz, S. and Johnson, N. (1992). "Distribution and Inferential Properties of Process Capability Indices," *Journal of Quality Technology*, 24, 216-231.

(11) Rodriguez, R.N. (1992). "Recent Developments in Process Capability Analysis," *Journal of Quality Technology*, 24, 176-187.

(12) Serfling, R.J. (1980). "Approximation Theorems of Mathematical Statistics," John Wiley and Sons, New York, 1-125.

(13) Spiring, F.A. (1991). "The $C_{pm}$ Index," *Quality Progress*, 24 (2), 57-61.

(14) Vännman, K. (1995a). "A Unified Approach to Capability Indices," *Statistica Sinica*, 5, 805-820.

(15) Vännman, K. and Kotz, S. (1995b). "A Superstructure of Capability Indices-Distributional Properties and Implications," *Scandinavian Journal of Statistics*), 22 (4), 477-491.

(16) Vännman, K. and Kotz, S. (1995c). "A Superstructure ofCapability Indices-Asymptotics and Its Implications," *International Jouranal of Reliability, Quality and Safety Engineering*, 2, 343-360.

86年11月10日 收稿
86年11月28日 修正
86年12月12日 接受

# 非常態製程下范氏製程能力指標估計式之漸近分佈

陳思勉
輔仁大學數學系

## 摘　要

Vännman（1995a）對常態製程提出一製程能力指標族 $C_p(a,b)$，並討論其估計式及相關之統計性質。本文則針對實務經常遇到的非常態製程討論該製程能力指標族之估計式的大樣本漸近分配。

# A Functional Approach to Finite Volume-Finite Difference Method

DANIEL LEE
*National Center for High-Performance Computing,*
*P. O. Box 19-136, Hsinchu, Taiwan, R.O.C.*
MULDER YU
*Department of Mathematics*
*Fu Jen University*
*Taipei, Taiwan 242, R.O.C.*

## ABSTRACT

A fourth order discretization to definite integrals frequently arising in the study of finite volume method was proposed in [6], in which analytic approach was adopted, while reduction to surface integral approach was taken in [4]. We present in the current paper a functional approach to this method, based on polynomial fitting. Application to a pure convection equation is discussed in some details, with general Crank-Nicolson (CN) approach. Stability result is estabilished and reveals that the scheme is unconditionally stable for CN parameter in the range 1/2 to 1, and of second order in time for CN parameter equal to 1/2.

***Key Words***: Finite Volume Method, Polynomial Fitting, Stability.

## INTRODUCTION

In the study of *finite volume method*, we are concerned with the equation

$$U_t + F_x + G_y = Q, \tag{1}$$

and its associated integral form

$$\frac{\partial}{\partial t}\iint_{\Omega_{i,j}} U dxdy + \iint_{\Omega_{i,j}} F_x dxdy + \iint_{\Omega_{i,j}} G_y dxdy = \iint_{\Omega_{i,j}} Q dxdy \tag{2}$$

Here $U$ is a scalar or vector function in temporal and spatial variables. The flux quantities, scalar or vector, depend on the spatial variables and possibly also on spatial derivatives. The control volumes form a partition of the computational space. There exist various methods in the construction of control volumes in the physical space [1, 3, 5, 7, 8]. In the current paper, we will consider only the *cell-center* type control volumes shown as in figures 1 and 2 that each grid point $x_i$ (or $(x_i, y_j)$) is at the center of the corresponding control volume $\Omega_i$ (or $\Omega_{i,j}$). Uniform mesh size is assumed. We refer to [1, 3, 7, 8] for details on the background and terminologies.

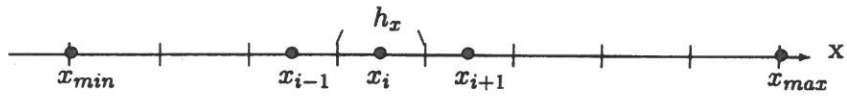



**图 1 one dimensional cell-center type control volumes.**

Some discussions in accuracy and stability, of a second order method, can be found in [8]. We will indicate a neat proof of these while we investigate a fourth order method later.

We derive in section 2 a functional approach to the fourth order discretization method [4, 6], via polynomial fitting, and then develop in section 3 the finite volume methods of second and fourth order in spatial variables. We also carry out in some details for the second order and fourth order approximations on a simple test example, to demonstrate the different approximation power. The fourth order method is further applied, in section 4, to a convection problem with Crank-Nicolson type approach. Theoretical result concerning stability is estabilished. Discussions and possible extensions are given in section 5.

## A FUNCTIONAL APPROACH TO DEFINITE INTEGRALS

The following two propositions are proved in [4, 6].

**Proposition 1** If $f \in C^4(\Omega)$ *with* $\Omega$ *a generalized rectangular domain, then*

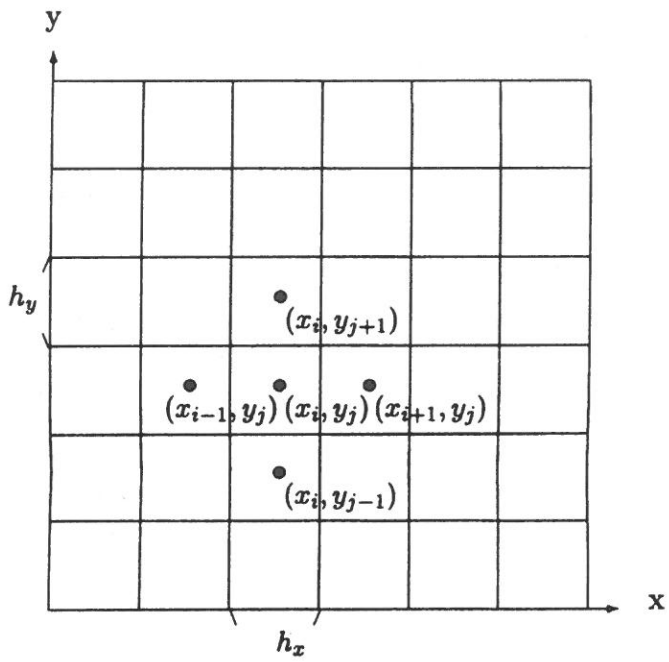



**圖 2 two dimensional cell-center type control volumes.**

$$\int_{\Omega} f(x)dx = |\Omega|\left(f(x_0) + \frac{1}{24}\sum_{i=1}^{n} h_{x_i}^2 f_{x_i x_i}\right) + O((h_{x_1} + \cdots\cdots + h_{x_n})^4) . \tag{3}$$

For three dimensional case this yields

$$\int_{z-\delta_z}^{z+\delta_z}\int_{y-\delta_y}^{y+\delta_y}\int_{x-\delta_x}^{x+\delta_x} f(x,y,z)dxdydz$$
$$= h_x h_y h_z\left( f(x_0, y_0, z_0) + f_{xx}\frac{h_x^2}{24} + f_{yy}\frac{h_y^2}{24} + f_{zz}\frac{h_z^2}{24} + O(|h|^4)\right). \tag{4}$$

Here and in the sequel,

$$h_x = 2\delta_x, \qquad h_y = 2\delta_y, \qquad h_z = 2\delta_z, \qquad |h| = (h_x + \cdots).$$

**Proposition 2** *If $f \in C^4(\Omega)$ with $\Omega$ a generalized rectangular domain, then for one-dimentional case we have*

$$\int_{\Omega} f(x)\, dx$$

$$= |\Omega|\left[f_i + \frac{1}{24}\left(f_{i+1} - 2f_i + f_{i-1} + O((h_x)^4)\right)\right]$$

$$= \frac{h_x}{24}[22f_i + f_{i+1} + f_{i-1} + O(|h|^4)], \tag{5}$$

*for two-dimensional case we have*

$$\int_\Omega f(x)\,dx$$

$$= |\Omega|[f_{i,j} + \frac{1}{24}(f_{i+1,j} - 2f_{i,j} + f_{i-1,j} + f_{i,j+1} - 2f_{i,j} + f_{i,j-1} + O((h_x + h_y)^4))]$$

$$= \frac{h_x h_y}{24}[20f_{i,j} + f_{i+1,j} + f_{i-1,j} + f_{i,j+1} + f_{i,j-1} + O(|h|^4)], \tag{6}$$

for three-dimensional case we have

$$\int_\Omega f(x)\,dx$$

$$= |\Omega|[f_{i,j,k} + \frac{1}{24}(f_{i+1,j,k} - 2f_{i,j,k} + f_{i-1,j,k} + f_{i,j+1,k} - 2f_{i,j,k} + f_{i,j-1,k} + f_{i,j,k+1} - 2f_{i,j,k}$$
$$+ f_{i,j,k-1} + O((h_x + h_y + h_z)^4))]$$

$$= \frac{h_x h_y h_z}{24}[18f_{i,j,k} + f_{i+1,j,k} + f_{i-1,j,k} + f_{i,j+1,k} + f_{i,j-1,k} + f_{i,j,k+1} + f_{i,j,k-1} + O(|h|^4)]. \tag{7}$$

In deriving the above, analytic approach was adopted in [6] while reduction to surface integral was taken in [4]. We give below a functional approach, based on polynomial fitting, with complete argument for three-dimensional case. Similar results also hold for one − and two − dimensional cases, respectively. Compared with the other two approaches, the current argument, although elementary, turns out to be the most appropriate one for extensions to nonuniform grids [2].

For convenience, we choose to work with the following formulation.

**Proposition 3** *If $f \in C^4(\Omega)$ with $\Omega$ a generalized rectangular domain, then*

$$\iiint fdxdydz = h_x h_y h_z(a_p f_{i,j,k} + a_w f_{i-1,j,k} + a_e f_{i+1,j,k} + a_s f_{i,j-1,k} + a_n f_{i,j+1,k}$$
$$+ a_b f_{i,j,k-1} + a_t f_{i,j,k+1}) + O((h_x + h_y + h_z)^4) \tag{8}$$

*with the constants $a_p$, $a_w$, $a_e$, $a_s$, $a_n$, $a_b$, $a_t$ depending only on the dimension.*

We derive firstly necessary conditions on the constants via a few test functions.

(1) The case $f = 1$ obviously implies

$$a_p + a_w + a_e + a_s + a_n + a_t + a_b = 1.$$

(2) For the case $f = x$, we calculate the left hand side of equation (8) as

$$\iiint_{x_i-\delta_x}^{x_i+\delta_x} x\,dx\,dy\,dz = \iint \frac{x^2}{2}\Big|_{x_i-\delta_x}^{x_i+\delta_x} dy\,dz = x_i h_x h_y h_z,$$

and obtain for the right hand side the following

$$\begin{aligned} & h_x h_y h_z(a_p x_i + a_w x_{i-1} + a_e x_{i+1} + a_s x_i + a_n x_i + a_t x_i + a_b x_i) \\ &= h_x h_y h_z(x_i(a_p + a_w + a_e + a_s + a_n + a_t + a_b) + h_x(a_e - a_w)) \\ &= h_x h_y h_z(x_i + h_x(a_e - a_w)). \end{aligned}$$

We conclude that $a_e = a_w$. Similarly, the cases $f = y$ and $f = z$ lead to $a_n = a_s$ and $a_t = a_b$ respectively.

(3) With the quadratic monomial $f = x^2$, we have

$$\begin{aligned} \iiint_{x_i-\delta_x}^{x_i+\delta_x} x^2\,dx\,dy\,dz &= \iint \frac{(x_i+\delta_x)^3 - (x_i-\delta_x)^3}{3} dy\,dz = \iint \frac{2(3x_i^2\delta_x + \delta S_x^3)}{3} dy\,dz \\ &= h_x h_y h_z\left(x_i^2 + \frac{h_x^2}{12}\right) \end{aligned}$$

and

$$\begin{aligned} & h_x h_y h_z(a_p x_i^2 + a_w(x_i - h_x)^2 + a_e(x_i + hx)^2 + a_s x_i^2 + a_n x_i^2 + a_b x_i^2 + a_t x_i^2)) \\ &= h_x h_y h_z(x_i^2(a_p + a_w + a_e + a_s + a_n + a_b + a_t) + x_i(-2h_x a_w + 2h_x a_e) \\ &\quad + h_x^2(a_w + a_e)) \\ &= h_x h_y h_z(x_i^2 + 0 + h_x^2(a_w + a_e)). \end{aligned}$$

These imply $a_w + a_e = 1/12$ and, therefore, $a_w = a_e = 1/2$. Similar argument with $f = y^2$ and $f = z^2$ then imply $a_s = a_n = 1/24$ and $a_b = a_t = 1/24$. We then end with

$$a_p = \frac{22}{24}, \frac{20}{24}, \frac{18}{24}$$

for one-, two- and three-dimensional cases, respectively. Actually, the above necessary conditions are also sufficient to make the equations (3), (4), (5) a fourth order discretization. To show this, we need to check out a few more cases.

In view of the identity $4xy = (x + y)^2 - (x - y)^2$ and the fact that both the integral and point evaluation functionals are linear in $f$, we expect the quadratic polynomial $f = xy$ to contribute nothing new. Actually, this can be justified directly. We check the argument below for two dimension only.

With $f = xy$, we calculate the left hand side of equation (8) as

$$\int_{y_j-\delta_y}^{y_j+\delta_y}\int_{x_i-\delta_x}^{x_i+\delta_x} xy\,dx\,dy = \int \frac{x^2}{2}\Big|_{x_i-\delta_x}^{x_i+\delta_x} y\,dy = \int_{y_j-\delta_y}^{y_j+\delta_y} y\,\frac{4x_i\delta_x}{2}\,dy$$
$$= (2\delta_y y_j)(2\delta_x x_i) = h_x h_y (x_i y_j) = h_x h_y f_{i,j},$$

and also the right hand side as

$$h_x h_y \frac{1}{24}(20(x_i y_j) + (x_{i+1} y_j) + (x_{i-1} y_j) + (x_i y_{j+1}) + (x_i y_{j-1}))$$
$$= h_x h_y \frac{1}{24}(20\, x_i y_j + y_j(x_{i-1} + x_{i+1}) + x_i(y_{j-1} + y_{j+1}))$$
$$= h_x h_y \frac{1}{24}(20\, x_i y_j + 2x_i y_j + 2x_i y_j) = h_x h_y x_i y_j.$$

This proves the current case.

It remains to check for cubic monomials.

(4) $f = x^3$ in one－dimensional case: The left hand side of equation (8) turns into

$$\int_{x_i-\delta_x}^{x_i+\delta_x} x^3\,dx = \frac{x^4}{4}\Big|_{x_i-\delta_x}^{x_i+\delta_x}$$
$$= \frac{1}{4}(x_i^4 + 4x_i^3\delta_x + 6x_i^2\delta_x^2 + 4x_i\delta_x^3 + \delta_x^4 - x_i^4 + 4x_i^3\delta_x - 6x_i^2\delta_x^2$$
$$+ 4x_i\delta_x^3 - \delta_x^4)$$
$$= \frac{1}{4}(8x_i^3\delta_x + 8x_i\delta_x^3) = h_x(x_i^3 + \frac{x_i h_x^2}{4}),$$

and the right hand side yields

$$h_x(\frac{22}{24}x_i^3 + \frac{1}{24}(x_i - h_x)^3 + \frac{1}{24}(x_i + h_x)^3)$$
$$= h_x(\frac{22}{24}x_i^3 + \frac{1}{24}(2x_i^3 + 6x_i h_x^2)) = h_x(x_i^3 + \frac{x_i h_x^2}{4}).$$

These establish the proposition for the case.

$f = x^3$ in two-dimensional case: We calculate for the left hand side

$$\int_{y_j-\delta_y}^{y_j+\delta_y}\int_{x_i-\delta_x}^{x_i+\delta_x} x^3\,dxdy = \int_{y_j-\delta_y}^{y_j+\delta_y}(h_x(x_i^3+\frac{xh_x^2}{4}))dy = h_xh_y(x_i^3+\frac{xh_x^2}{4}),$$

and for the right hand side,

$$h_xh_y\Big(\frac{20}{24}x_i^3+\frac{(x_i-h_x)^3+(x_i+h_x)^3+x_i^3+x_i^3}{24}\Big)$$
$$= h_xh_y\Big(\frac{20}{24}x_i^3+\frac{4x_i^3+6xh_x^2}{24}\Big) = h_xh_y(x_i^3+\frac{xh_x^2}{4}).$$

These justify equation (8).

Finally, we consider the cases $f = x^2y$ in two dimension and $f = xyz$ in three dimension. (5) $f = x^2y$ in two-dimensional case: The left hand side reduces to

$$\int_{y_j-\delta_y}^{y_j+\delta_y}\int_{x_i-\delta_x}^{x_i+\delta_x} x^2\,y\,dxdy = \int_{y_j-\delta_y}^{y_j+\delta_y}(h_x(x_i^2+\frac{h_x^2}{12}))y\,dy = h_xh_y(x_i^2+\frac{h_x^2}{12})y_j$$

and the right hand side yields

$$\frac{h_xh_y}{24}(20\,x_i^2y_j+(x_i+h_x)^2\,y_j+(x_i-h_x)^2\,y_j+x_i^2(y_j+h_y)+x_i^2(y_j-h_y))$$
$$= \frac{h_xh_y}{24}(20x_i^2y_j+2\,x_i^2y_j+2\,h_x^2y_j+2\,x_i^2y_j) = h_xh_y(x_i^2+\frac{h_x^2}{12})y_j.$$

The desired equation (8) then follows.

(6) $f = xyz$ in three－dimensional case: We simplify the left hand side of equation (8) to get

$$\int_{z_k-\delta_z}^{z_k+\delta_z}\int_{y_j-\delta_y}^{y_j+\delta_y}\int_{x_i-\delta_x}^{x_i+\delta_x} xyzdxdydz = (x_ih_x)(y_jh_y)(z_kh_z) = h_xh_yh_z(x_iy_jz_k),$$

and note that the right hand side yields

$$\frac{h_xh_yh_z}{24}(18\,x_iy_jz_k+y_jz_k(x_i+h_x+x_i-h_x)+x_iz_k(y_j+h_y+y_j-h_y)$$
$$+x_iy_j(z_k+h_z+z_k-h_z))$$
$$= \frac{h_xh_yh_z}{24}(18x_iy_jz_k+2\,x_iy_jz_k+2\,x_iy_jz_k+2\,x_iy_jz_k) = h_xh_yh_z(x_iy_jz_k).$$

We conclude therefore the discretization in Proposition 3 is exact for polynomials of degree less than four.

## A FOURTH ORDER FINITE VOLUME METHOD

By application of equation (4), we obtain the following two discretization schemes for equations (1) and (2).

**FV2 Discretization:**

$$(U_t)_{i,j} + (F_x)_{i,j} + (G_y)_{i,j} = Q_{i,j} + O((h_x + h_y)^2). \tag{9}$$

**FV4 Discretization:**

$$\begin{aligned}
&20(U_t)_{i,j} + (U_t)_{i-1,j} + (U_t)_{i+1,j} + (U_t)_{i,j-1} + (U_t)_{i,j+1} + \\
&20(F_x)_{i,j} + (F_x)_{i-1,j} + (F_x)_{i+1,j} + (F_x)_{i,j-1} + (F_x)_{i,j+1} + \\
&20(G_x)_{i,j} + (G_x)_{i-1,j} + (G_x)_{i+1,j} + (G_x)_{i,j-1} + (G_x)_{i,j+1} \\
&= 20(Q_x)_{i,j} + (Q_x)_{i-1,j} + (Q_x)_{i+1,j} + (Q_x)_{i,j-1} + (Q_x)_{i,j+1} + O((h_x + h_y)^4).
\end{aligned}$$

To demonstrate the different approximation power of the two discretizations, we consider the following simple test function

$$u(x,y) = x^2 + 2xy + y^2,$$

and its integral

$$\iint_{\Omega_{ij}} u(x,y)dxdy = \iint_{\Omega_{ij}} (x^2 + 2xy + y^2)dxdy$$

on a single control volume $\Omega_{i,j} = [x_i - (h_x/2), x_i + (h_x/2)] \times [y_j - (h_y/2), y_j + (h_y/2)]$.

We set $(x,y) = (x_i + \xi,\ y_j + \eta)$ and calculate the above definite integral as follows.

$$\begin{aligned}
&\iint_{\Omega_{ij}} (x^2 + 2xy + y^2)dxdy \\
&= \int_{\frac{-h_y}{2}}^{\frac{h_y}{2}} \int_{\frac{-h_x}{2}}^{\frac{h_x}{2}} ((x_i + \xi)^2 + 2(x_i + \xi)(y_j + \eta) + (y_j + \eta)^2)d\xi d\eta \\
&= \int_{\frac{-h_y}{2}}^{\frac{h_y}{2}} \left( \frac{(x_i + \frac{h_x}{2})^3 - (x_i - \frac{h_x}{2})^3}{3} \right) d\eta
\end{aligned}$$

$$+2\Big(\frac{(x_i+\frac{h_x}{2})^2-(x_i-\frac{h_x}{2})^2}{2}\Big)\Big(\frac{(y_j+\frac{h_y}{2})^2-(y_j-\frac{h_y}{2})^2}{2}\Big)$$

$$+\int_{\frac{-h_x}{2}}^{\frac{h_x}{2}}\Big(\frac{(y_j+\frac{h_y}{2})^3-(y_j-\frac{h_y}{2})^3}{3}\Big)d\eta$$

$$=\frac{2h_y}{3}\Big((3x_i^2\frac{h_x}{2}+(\frac{h_x}{2})^3\Big)+2\Big(2x_i\frac{h_x}{2}\Big)\Big(2y_j\frac{h_y}{2}\Big)+\frac{2h_x}{3}\Big(3y_j^2\frac{h_y}{2}+(\frac{h_y}{2})^3\Big)$$

$$=x_i^2h_xh_y+y_j^2h_xh_y+\frac{h_x^3h_y}{12}+\frac{h_xh_y^3}{12}+2x_iy_jh_xh_y$$

$$=h_xh_y((x_i^2+2x_iy_j+y_j^2)+\frac{1}{12}(h_x^2+h_y^2)).$$

While the FV2 discretization of the definite integral yields

$$h_xh_y(x^2+2xy+y^2)_{i,j}=h_xh_y(x_i^2+2x_iy_j+y_j^2),$$

the application of the FV4 discretization, together with the fact that

$$u_x=2x+2y,\ u_y=2y+2x,\ u_{xx}=2,\ u_{yy}=2$$

then results in

$$h_xh_y\Big(x_i^2+2x_iy_j+y_j^2+\frac{1}{24}(2h_x^2+2h_y^2)\Big)=h_xh_y\Big(x_i^2+2x_iy_j+y_j^2+\frac{1}{12}(h_x^2+h_y^2)\Big).$$

Comparing the last three equations, it is clearly seen that the fourth-order discretization is exact for the current integrand, as it should be, and the second-order discretization produces approximation error of order 2. Further application to a pure convection equation is given next.

## LINEAR STABILITY

From the derivation in previous sections, it is obvious that the FV2 and FV4 methods, as in their simplest form, are, in general, both first order in time, second order and fourth order in space, respectively. To investigate the Von Neumann stability, we consider the *Pure Convection Equation*

$$u_t+a\,u_x=0$$

and apply general Crank-Nilcoson version of the FV4 method to its integral form

$$\int u_t + \int a\, u_x = 0.$$

We begin with a consistent finite difference approximant

$$\int_{\Omega_i} \frac{u_i^{n+1} - u_i^n}{h_t} + \int_{\Omega_i} a\, \frac{-u_{i+2}^n + 8\,u_{i+1}^n - 8u_{i-1}^n + u_{i-2}^n}{12\,h_x} = 0$$

and apply the FV4 scheme to obtain

**FV4-CN Method:**

$$\begin{aligned}
&\frac{22u_i^{n+1} + u_{i+1}^{n+1} + u_{i-1}^{n+1}}{24} - \frac{22u_i^n + u_{i+1}^n + u_{i-1}^n}{24} \\
&+ \frac{ah_t}{12\,h_x}\frac{1}{24}\,(\alpha((-22u_{i+2}^{n+1} - u_{i+3}^{n+1} - u_{i+1}^{n+1}) + (22u_{i-2}^{n+1} + u_{i-1}^{n+1} + u_{i-3}^{n+1}) + 8(22\,u_{i+1}^{n+1} + u_{i+2}^{n+1} \\
&+ u_i^{n+1}) \\
&- 8(22u_{i-1}^{n+1} + u_i^{n+1} + u_{i-2}^{n+1})) + (1-\alpha)\,((-22u_{i+2}^n - u_{i+3}^n - u_{i+1}^n) \\
&+ (22u_{i-2}^n + u_{i-1}^n + u_{i-3}^n) + 8(22u_{i+1}^n + u_{i+2}^n + u_i^n) - 8(22u_{i-1}^n + u_i^n + u_{i-2}^n))) \\
&= 0.
\end{aligned}$$

**Proposition 4** *If the CN parameter $\alpha$ satisfies $1/2 \leq \alpha \leq 1$, then the FV4-CN method is unconditionally stable.*

The symbol, or amplification factor, is calculated as follows

$$\begin{aligned}
&g(22 + e^{i\theta} + e^{-i\theta}) - (22 + e^{i\theta} + e^{-i\theta}) \\
&+ \frac{ah_t}{12h_x}(\alpha g(-(22 + e^{i\theta} + e^{-i\theta})\,e^{-i2\theta} + (22 + e^{i\theta} + e^{-i\theta})e^{i2\theta} + 8(22 + e^{i\theta} + e^{-i\theta})e^{i\theta} \\
&- 8(22 + e^{i\theta} + e^{-i\theta})e^{-i\theta}) \\
&+ (1-\alpha)(-(22 + e^{i\theta} + e^{-i\theta})\,e^{-i2\theta} + (22 + e^{i\theta} + e^{-i\theta})\,e^{i2\theta} + 8(22 + e^{i\theta} + e^{-i\theta})\,e^{i\theta} \\
&- 8(22 + e^{i\theta} + e^{-i\theta})\,e^{-i\theta})) \\
&= 0.
\end{aligned}$$

This yields

$$\begin{aligned}
&g - 1 + \frac{a\,h_t}{12h_x}(\alpha g(-e^{i2\theta} + e^{-i2\theta} + 8\,e^{i\theta} - 8e^{-i\theta} \\
&+ (1-\alpha)\,(-e^{i2\theta}e^{-i2\theta} + 8e^{i\theta} - 8e^{-i\theta}))
\end{aligned}$$

$= 0.$

We obtain therefore the symbol

$$g = \frac{1 - \frac{ah_t}{12h_x}(1-\alpha)(16sin\theta - 2sin2\theta)i}{1 - \frac{ah_t}{12h_x}\alpha(16sin\theta - 2sin2\theta)i},$$

and consequently,

$$|g|^2 = \frac{\left|1 - \frac{ah_t}{12h_x}(1-\alpha)(16sin\theta - 2sin2\theta)i\right|^2}{\left|1 - \frac{ah_t}{12h_x}\alpha(16sin\theta - 2sin2\theta)i\right|^2}.$$

The expecting result $|g| \leq 1$ now follows under the assumption $\frac{1}{2} \leq \alpha \leq 1$.

We note more stability and convergence results for a general convection-diffusion equation in one and two dimensional spaces are provided in [6].

## CONCLUSION

It is presented in this paper an elementary argument to derive an analytic approximation to definite integrals frequently arising in the application of finite volume method. Further difference formulae then yield discrete algebraic system for each individual problem. Preliminary study [2] show that the current approach is easier to extend to non-uniform grid case than alternative approaches taken in [4, 6]. Numerical experiment [9], with a two-dimensional nonlinear convection-diffusion problem, confirms the orders in accuracy and convergence of the method.

## ACKNOWLEDGEMENT

The authors want to thank the referee for suggestions toward improvment of this paper. Very sincere gratitudes are expressed here.

## REFERENCES

(1) J. D. Anderson, Jr., *Computational Fluid Dynamics: The Basics With Applications*, *McGraw-Hill*, *Inc.*, 1995.

(2) C. W. Chen and Daniel Lee, *A Fourth Order Finite Volume-Finite Difference Method for Nonuniform Grids (in preparation)*.

(3) C. Hirsch, *Numerical Computation of Internal and External Flows*, *Vol*. 1, 2 : *Fundamentals of Numerical Discretization*, *John Wiley and Sons Ltd*., 1988.

(4) K. C. Jea, Daniel Lee and Mulder Yu, *A Fourth Order Finite Volume-Finite Difference Method*, *NSC Project Report*, *Taiwan*, *R*. *O*. *C*., *January* 1998.

(5) R. D. Lazarov, I. D. Mishev and P. S. Vassilevski, *Finite Volume Methods for Convection-Diffusion Problems*, *SIAM J*. *Numer*. *Anal*., *Vol*. 33, *No*. 1, *pp*. 31-55, *February* 1996.

(6) Daniel Lee, *A New Approach to Finite Volume - Finite Difference Methods*, *NCHC Technical Report*, *Taiwan*, *R*. *O*. *C*., *August* 1997.

(7) K. W. Morton, *Numerical Solution of Convection - Diffusion Problems*, *Chapman and Hall*, 1996.

(8) J. W. Thomas, *Numerical Partial Differential Equations : Finite Difference Methods*, *Springer-Verlag*, *New York*, 1995.

(9) Mulder Yu, *A Study on Finite Volume Methods*, *Master Thesis*, *Department of Mathematics*, *Fu-Jen University*, *June* 1997.

86年11月12日 收稿
86年12月 6日 修正
86年12月23日 接受

# 從泛函的觀點探看有限體積-有限差分法

李 天 佑
國家高速電腦中心
游 輝 宏
輔仁大學數學研究所

## 摘 要

參考［6］針對有限體積計算方法常處理之定積分提出了一個四階離散化之作法，該作法是以解析方式探討。參考文［4］則以面積分之處理方式探討。

本文將以泛函方式發展多項式配湊法繼續探討，並應用至對流方程式，以 Crank-Nicolson 方法發展離散方程。由此並證明當 CN 參數介於 1/2 與 1 之間時，此一數值方法爲無條件穩定；當 CN 參數等於 1/2 時，此一方法對時間變數爲二階準確。

**關鍵詞：** 有限體積方法，多項式配湊法，穩定性。