師大數學

19

序

系主任

本系成立於民國三十五年八月，首由張儒林教授擔任系主任，當時招收數學系及四年制專修科各一班，同學三十餘人，教授三位，助教一位，嗣由管公度、潘璞、李新民、康洪元、范傳坡、常法徽各教授輪掌系務，歷經各任主任與教授們之群策群力，始具今日之規模。三十九年來本系之畢業系友，已逾貳仟陸佰餘人，多各有成就；其中獲得碩士學位叁佰餘人，具博士學位者逾一百五十人，或執教於高等學府，或繼續高深研究，餘多服務於中等教育界，平時諄諄教學，誨人不倦，敦品篤行，懷抱熱忱，此實本校優良風氣之所致。

現本系有教師四十七人，學生方面日間部有十二班，夜間部兩班，共有同學約伍百人，圖書逾兩萬册，雜誌百餘種，微型電腦四十八部，電算室二間，自六十四年夏遷於現址後，環境煥新，出國學成系友或返校互相砥勵，研究風氣已大弧度地提高；今日數學系之師生孜孜不息，無不爲美好遠景而奮發。

近來科學發展甚速，對數學之需要甚切，本系之任務更日益增大，除肩負數學教育之發展及中等數學教育之輔導外，亦兼具數學學術及數學教育研究之重要任務。爲增强研究風尚，本系於十八年前創辦師大數學年刊，以供師生發表教學及研究心得。切磋琢磨，提高學習及研究興趣，屢經負責同學之辛勞耕耘及師生系友之共同支持。漸茲茁壯，今後請更不吝珠璣，源源賜稿，則本刊必前程似錦，散發絢爛光芒，於此，表謝忱，更盼我師生系友善珍此片園地。

顏啟麟 謹識

七十四年五月

目　錄

1.微積分的沒落——
離散數學的興起

指導老師：洪萬生老師
一　　丙：柯遜富

前言：

本文是翻譯The Decline of Calculus——The Rise of Discrete Mathematics（Anthony Ralston）取材自Mathematics Tomorrow p.213（凡異出版社，1981）。

由於這是本人第一次從事翻譯、投稿，再加上數學知識並不充分，因此對於句意的表達，文章之編排不免有疏漏及錯誤之處，尚祈見諒，並請諸先進不吝指正。

專門性的來講，分析學是數學中最成功並最精巧的部份。

——John Von Neumann〔1951〕

在即將來臨的數十年或世紀中，關於計算機有一個簡單而基本的事實，那就是計算機所影響的，與其說是數學中已經知道的，倒不如說其中被認爲重要的東西。這就是它的有限性。——Wallace Givens〔1966〕

微積分是人類智力最偉大的勝利品之一。單單爲了這個理由，每個受教育的人就應該有一點有關的知識。此外；當你想到古典分析學（其基礎爲微積分）的全套知識及實用的勝利品時，對於微積分在這麼長的時間一直爲所有學院數學課程之基礎，就不會感到太驚訝了！但也因此，當讀者知道這篇文章主討論微積分在學院數學課程裏的位置，已經完全成熟到即將有所改變，以及多少有點衰退時，就很有理由驚訝了！

假如你和我一樣，相信數學研究和其它智力活動一樣需要高度創造力及想像力時，那麼無論如何，當那些好的數學家從事他們的事業時，一定是一

群最不保守的人們（在這裏“保守”意味著執著於已建立的思想模式）。眞的，成功的數學研究家一定要個準備去接受——或最少去考慮——最狂妄新觀念的急進分子。但他們並非只在直接的事業活動之中才是急進分子，至少在他們觀察中小學教育時，他們也有許多是教育的急進分子。本文的目的不在討論「新數學」但不管你把它看做什麼，這些觀念都是激進的。無論如何，相當奇怪的是當數學家來到自己的天地——大學的數學課程時，他們便成了最保守的代表。

在冒點過度簡化的危險下，我們可以這麼說：這個世紀裏唯一的重大改變就是：直到二次大戰以後，還是學院數學課程大一新生標準科目的「學院代數」及「三角學」。已經被兩年的微積分及解析幾何所取代而降至高中課程，然後後者又爲微積分及線型代數所取代。二次大戰以前大學數學系的學生直到二年級才涉獵微積分。在1960年代，在Dartmouth有一些值得注意的短暫實驗，即在連續的頭兩年裏包含了一學期的「有限數學」，但在絕大部分的情況下，這些課程現在已被線型代數所取代。最近，正當爲了新鮮人的數學預備知識似乎已經退步時，學院代數及三角學又回到學院數學裏復辟了。但這些變數並不影響基本的觀點；在本世紀，前兩年的學院數學裏已有了一些變異。

我並不主張事物應有很大的差異。小學、中學或大學新生水準之任意學科的主要預備課程中，是不應該有快速的改變或不連貫。知識上的根本改變是很少見的；更少的是這些改變蘊涵著在基本教育水準相當多的變化。此外，我們對於區別好、壞教育的方法之了解是如此的少，有用的評量我們對教育所做一切努力是如此的少，使得我們應該以高度的懷疑主義來觀察那些明顯的改變，並僅僅爲了最强烈的理由而被迫從事那些改變。

然而，我仍然建議如此的革命是需要的，原因何在呢？在過去三十年當中，數字型電子計算機之發明及發展，或許是自從印刷術的發明以來，科學及技術上最重要的發展。無論如何，這發展不僅在人類生活及社會結構上具有深遠的影響，而且也將會——此亦爲這裏的主要觀點——在科學家研究，特別是他們所用的數學上有重大的影響。（我在此所强調的，並不是說微積

分及古典分析學將不再繼續有大進展，而只是說它們在數學裏的支配位置及應用即將要受到挑戰。）

這革命的內容是（或應該是）什麼呢？沒有東西比得上「離散分析」及「古典連續分析」某種（至少）對等更少的了！對於離散分析，我的意思是包括那些數學的分支，它們主要或完全注重於離散的對象，包括：組合理論、圖形理論、抽象代數、線型代數、數論和離散機率。這對等性至少由一種以離散數學爲基礎的一、二年級序列的預備，做爲取代標準微積分的序列，便是以顯示它自己。（經由如此的序列，三年級數學的標準課程，將會是微積分了。）

這個建議的基本動機是：對許多直接處理電子計算機的人（包括至少大部分的電子計算機科學家，以及社會、行爲和管理的科學家）而言，數學中最重要的似乎不是微積分而是離散數學的領域。在這個前提之下，無論如何，我相信數學研究在離散數學這個方向的趨勢是强烈的增加；一個主張雖然並未加以確切證實，但能由數學研究出版品的趨向研究予以支持。除了任何憑經驗的證據之外，這個命題可由以下的觀察來支持：相當大程度來說，數學泉源一直總是數學的應用。今日，一般的電子計算機及特別的電子計算機科學家，比起任何其它科學及工業的領域，對應用數學產生了更大的需要量，而且還以一個相當快的速度增加。由於經由電子計算機及電子計算機科學家所產生的數學問題，壓倒性地化連續數學工具更需要離散，所以在離散數學研究快速的成長，就不會令人驚訝的了。以下，我將試著提出一些觀點，解釋爲何離散數學在與電子計算機相關的數學裏，扮演如此重要的角色。

首先，讓我提示什麼是至少在目前一個注重離散數學的相當不重要之理由。由於不顧他們所塑造的物理狀態是眞正的離散系統這一事實，使古典分析學獲得了不起的成就。例如：實際上一個個體的運動是它單一的、離散的微分子運動之集合體。無論如何，由於這些分子的總數太大了，以致於我們唯有將他們當成一個個體才有可能獲得一個有意義的結果。當然，由於如此的接近眞實，使得以這個方法去做而獲得的結果通常眞的都十分正確。有點諷刺的，把古典分析的公式應用在電子計算機上來做，眞正的計算只能由一

次離散這些公式（例如：用和來取代微分及積分的導數）來完成。（無論如何，這離散化，將比只考慮單一的微分子個體，產生更不詳細的模型）。因此，我們便有了這種情況，那就是：本質上的離散現象由連續函數所塑造而成，而因爲計算之目的，連續函數必須加以離散化。

爲此說來，在一開始就離散地處理離散問題不是能獲得更多的意義嗎？到目前爲止，這僅僅是說說而已。做這件事所要計算的狂野行動，仍然遠超過可利用的最大電子計算機之能力。但這不會永遠是眞的，當微處理機來臨時，將能策勵計算機網路的發展。假如那時候如此的計算變爲可能的話，那它們將能夠在眞實物理及實際的計算間提供一個連結，此連結在解釋及應用物理理論將是最有幫助的。但這時代尚未到來：物理實在的分析現在並不是强調離散分析重要性的迫不得已理由。

算則：

對於我想注目的離散數學之應用，就是它在算則分析的用途，這應用在計算一些數值或符號之值是個完備定義的程序。由此幾乎全部計算機的程式是算則的實現，能否回答以下二個基本問題就顯得極端重要：

1 是否一個程序卽爲一個算則？也就是說，它是否能很可靠地引導到一個想要的結果，或者（最壞的情況下）能給一個「它不能解答」的指標？

2 一個特別的算則——絕對的或關連到同一問題的其它可能算則——多有效力呢？

回答這兩個問題是算則分析的主要問題。

要說明伴隨算則分析的數學之問題及種類，讓我們考慮一個特別的問題（其本身並沒有太大內在的重要性）。

在Baltimore Hilton Inn（以下簡稱B.H.）的每個房間裏，不用一般的鎖，而用一個輸入四個數字組合的號碼盤。只要撥了任意組合的四位數字（依正確順序）鎖就會開了！假設一個賊在完全不知道這組合的情況下而想進入一個房間，自然撥得越少越好。雖然可能有40000個數字要撥（因爲有10^4種組合，每種撥4個數字），但經過適當的設計將能較少。例如：撥了abcd後再撥e，就能有效率地試bcde了。這個問題就是去決定試過所

有組合後所要撥的最少次數，並去發現一個有效率的算則產生一個最短長度之序列（熟悉圖形理論的讀者將能記起這如同Bruijn cycle 問題的再版）。

如同數學家所習慣的一樣，我們開始去推廣這個問題，假設有m個不同的符號（在B.H.中m＝10）當一個最短的序列，能包括所有由m所形成長度爲n（在B.H.中n＝4）的組合時，正是我們所想要的。一個低限是m^n+n-1。因爲有m^n個組合，產生第一個組合要用一串長度n，然後每加一個符號便最多能產生（m^n-1）的組合中之一個。

這個低限是否已完成了，或者必定包含一些重複的長度n之序列？這答案可由離散數學中重要一支——圖形理論中的一個衆所皆知的理論找到答案。實際上這個理論證明所有重複皆能避免，使得這最短長度正是我們所要找的低限。

現在，讓我們來考慮一個有m^{n-1}個點的圖。於此，每一個點分別被標爲長度爲（n－1）由m個不同符號所組成之m^{n-1}組合中之一個。圖1是這種圖的一部分，且假設m個符號爲0，1，2，……（m－1）。一個標爲$m_1 m_2 \cdots m_{n-1}$的點有m個邊從它發散出去分別到標爲$m_2 m_3 \cdots m_{n-1}\, i$

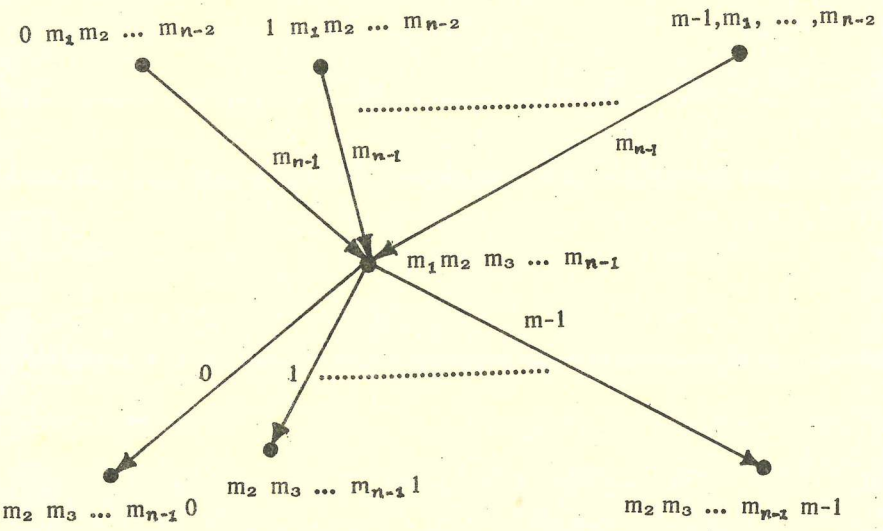


圖 1

的點（$i=0,1,2,\cdots\cdots,m-1$）每一個邊分別用i來標記。

此外每個點亦有附屬的m個邊，從$i\ m_1m_2\cdots\cdots m_{n-2}$（$i=0,1,2,\cdots\cdots m-1$）到點$m_1m_2\cdots\cdots m_{n-1}$，每一個邊皆記爲$m_{n-1}$。

這理論需要分析的敍述：如圖1這種圖（專門一點而言，一連結、直接的圖）從任何一點開始，都有一個路徑（即尤拉路徑）包括每一個邊一次且僅有一次。當然，前提是每個點都有m個邊發射出去及m個邊射進來。（這只是對流行兒童遊戲形式上的解。這遊戲便是沿著一幾何圖形而筆不離紙或走過同樣線段兩次。如圖1這種圖每一點皆有m個邊從它發出及射入，並且是相連的，故從任何一點開始，它一定是個「尤拉路徑」。

現在，假設我們從$m_1m_2\cdots\cdots m_{n-1}$開始，並從圖的這一點走尤拉路徑。每當我們走過一個邊，我們就在這一序列加入這個邊的記號。（如此，當從點$m_1m_2\cdots\cdots m_{n-1}$到點$m_2m_3\cdots\cdots m_{n-1}\ i$〔1〕，我們就加$i$到這一序列）當每一次走過一個邊時，這序列之最後n個便形成了任一之組合。此外，由於每一個點及邊的標記，故沒有一個組合能重複兩次以上。因爲有m^n個邊（有m^{n-1}個點，每一點有m個邊）而第一個出發的點含（n－1）個，所以，當我們走完尤拉路徑時，序列剛好（m^n+n-1）個，如以上所說，在m^n排列中並沒有重複。

用這一序列（m^n+n-1）存在的證明，我們便可以很容易的將產生這序列的算則公式化：

（Ⅰ）以n個0來展開這一序列，這好比從點（由（n－1）個0標記）開始走邊（標爲0）。這將形成一個環而走到你開始之點。

（Ⅱ）在每個階段加入一個非0之數，如果產生一個新的組合，便加此數，否則便加一個0。

雖然並不十分明顯，但不難證明這些步驟將直到你回到原來的點（m＋1）次，並且所有的邊都走過一遍，才會終止。（一個尤拉路徑以始點爲終點，稱爲尤拉循環。）

這個問題完了嗎？或許，如同一些數學家所認爲，我們已證得了一個存在的答案，並知道如何去找這答案，但一個並不以此滿足的計算機科學家將

會想知道能否以計算機來更有效率地完成這個算則。最大的問題在第二步，於此每一個可能的符號必定引發出廣泛的「檢查重複」以保證加下去這個符號將不會產生已有的組合。在B.H.中m＝10，n＝4，有10000個組合之記憶需要儲存使人們能在它們出現在序列時便能找出。此外，我們必須在每次加入一個數字後，查表9次，直到發現沒有重複。一般而言，一個有m^n的值之表，每次將要找（m－1）次。

有一個可能的方法去改良這空間效率（常叫做空間複雜性），就是去儲存10000連續的微片（二進位）以代替字，若這序列沒有重複則顯示0，若有則顯示1。但這將會在時間複雜性造成大量消費，因爲在典型相當小的微片中找尋比在用字作成的表找訊息還較沒效率。

我們所想要的是一個能使我們直接加入數字到序列中而不必檢查的建構性算則。如此的算則的確存在，但它們太過於複雜，故我們在此只能概述一個。

（Ⅰ）在一開始先寫一個單一數字（m－1）。

（Ⅱ）再增添（m－1）串，長爲n之序列：（m－1）（m－1）……（m－1）（m－2），每一個序列前（n－1）個數字都一樣，但最後一個則每次遞減1。（在B.H.中由（Ⅰ）、（Ⅱ）可得：

9　9998　9997　9996……9990）

（Ⅲ）再來添加$(m-1)^2$串長度爲n之序列：（m－1）（m－1）……（m－1）（m－2）（m－2），於此，每一個序列有（n－2）個（m－1）且後面緊跟著兩位由0，1，……（m－2）所任意組合之最大兩位數，但前者永遠大於後者（在B.H.中此產生了：

9988　9987……9980　9978……9970……9900）

（Ⅳ）現在由於複雜性已產生了，所以一定很模糊、混亂，我們繼續做長爲n之序列，在每個階段，有系統的引進更少及不同位置的（m－1），無論如何，每一個序列（長爲n）依賴它的前者是一個完備定義。

（Ⅴ）最後我們爲了（m－1）及n再增加相當的序列，我們以（m－1）代替m應用（Ⅰ）～（Ⅳ），再用（m－2）代替（m－1）做（Ⅰ）～（

Ⅳ)……。由於到了m＝1及任意n時剛好n個0，所以應用(Ⅰ)～(Ⅳ)(m－1)次就做完了這步驟。

(Ⅵ) 最後一步，我們在整個序列加入(m－1)這個數字即可。

這個算則有幾點值得注意：首先，由於這序列是依賴前者爲完備定義，故這算法並不需要記憶以前是否產生過這序列。並且由於這一序列由前者所產生之規則相當簡單，所以這規則無論從時間或空間的觀點而言，都相當有效率。

步驟(Ⅱ)～(Ⅳ)牽涉到迭代(iterative process)(亦即一個序列接一個序列而產生)，而此正是在離散數學及算法中許多問題的特色。步驟(Ⅴ)包含了一「循環」，它是一個有特別參數(m)問題的解，它依賴一連串比那參數較小的參數(m－1，m－2，……，1)。「循環」對數學家及計算機科學家而言，是一個可供利用，最有力的工具之一。

最後，雖然我們的討論並沒有顯示，或許相當驚訝的，竟然發生在算則的細節上：當n是質數比n不是質數時要求的相當簡單。

總而言之，這個問題，連同其解代表了牽涉到算則分析的問題，並且也代表了離散數學中各種領域(圖形理論、組合、甚至數論)和算則的分析及設計之間的交互影響，但一個例子不能證明此一事實。不過我堅信這個問題連同它設計及解析的算則是如此的廣泛，如此快速而愈形重要，並且遺留下的重要是如此的久，使得在離散數學的研究將會有快速的進步以提供技術及結果給需要的算則。於是我又回到了我在本文一開始的論點：假如尚未過時，離散數學將會在大學生的數學課程裏，至少被認爲如同古典解析之伙伴。而現在正是時候了!!

註釋：

註〔1〕：原文爲$m_2 m_3 \cdots\cdots m_{n-2}\ i$。

註〔2〕：(Ⅵ)爲本人所加，原文並沒有。讀者有興趣，可加以舉例，並證明這六個步驟是正確的。

參考書目：

(1)Daniel Greenspan; Discrete Models(Addison-Wesley, 1973)。

(2)Donald E. Knuth; The Art of Computer Programming V.1, 2, 3. (Addison-Wesley, 1968, 1969, 1973)。

(3)Anthony Ralston; 「Computer Science, Mathematics and the Undergraduate Curricula in Both」(American Mathematical Monthly)。

(4)Anthony Ralston; 「A New Memoryless Algorithm for De Bruijn Sequences」(Journal of Algorithms)。

本文承蒙洪萬生老師、費毓港助教指導，謹此致謝。

2. Baker's Transformation is Ergodic

指導老師：王惠中老師
譯　　者：費毓港

原文：P.R. Halmos　Lectures on Ergodic Theory
The Mathematical Society of Japan (1956)
第1，3，8節。

來由：Halmos 在1955年於芝加哥大學暑期課程的講義。

內容：Ergodic 理論的入門基本知識。

第1節　「簡介」　從Ergodic理論的歷史談到保測變換的來由。

第2節　「例子」　介紹一些保測變換的例子，引入Baker's transformation。

第8節　「遍歷性」　談保測變換的遍歷性，主要是是線性變換和緊緻群上的自同構，證明Baker's transfomation是遍歷的(ergodic)。

題目：是譯者自己定的。

翻譯目的：讓讀者能夠瞭解題目的那句話。

翻譯動機：王惠中老師灌輸一些ergodic 知識的時候，曾經說過Halmos的文筆很好，這本書也寫得不錯；而譯者想翻一些東西，「硬」的又不懂，只好挑比較簡單的，遂成此篇。

預備知識：要先知道下面一些名詞的意思：
集合、函數、連續、緊緻、群、測度、歐氏空間、Borel集、吃飯、走路、讀書、……。

翻譯時間：1985年春節期間。

心得：Halmos 這本書的資料在ergodic 理論上已經不算新的了（蠻舊的），然其含意深遠，娓娓道來，直像說故事一樣的，可讀性甚高。

譯者讀過（最少前面幾章）Halmos 老兄的好幾本書，諸如Measure Theory，Finite Dimensional Vector Space，Hilbert Space，Naive Set Theory 和這本Ergodic Theory，覺得前兩本頗爲「格律」化，後面兩本就比較有「味道」了。順帶一提的是Naive Set Theory 的序言裏有「Read it, absorb it, and forget it」這樣的「金句」，而Ergodic Theory 裏面連序言都省了，一開始就來一道「Apology」，可以說是互相輝映了。

附記：(1)Ergodic 理論起源自統計力學，由H. Poincare 引入，歷經G. D. Brikhoff，B. Koopman，A. Kolmogorov，Y. Sinai，D. Anosov，D. Ornstein 等大高手的努力後，已經變成純數學的一個新的分支。

(2)Dieudonne 在他的A Panorama of Pure Mathematics中也承認Halmos 在Ergodic 理論上有過實質的貢獻。

(3)感謝上天的幫忙，寒假期間天天下雨，使我「能夠」待在家裏，很快就完成這篇的翻譯了。（賊老天！）

簡 介

哥尼斯堡（Koenigsberg）的七橋問題引發起對拓樸學的研究；而遍歷理論的探討則來自於在統計力學裏的一些沈思。七橋問題在數學上的發展，得到一個關於圖形奇、偶頂點的定理；氣體粒子的問題經過數學的延伸，讓我們對保測變換（measure preserving transformations）的漸近行爲（asymptotic behavior）有深入一層的了解。無論是那一種情形，直接從那原始的動機所導出的結果，只佔整個廣濶理論的一小部份。可是，明白一個理論的歷史背景，畢竟有其眞確的價值，因此，我將會對統計力學中的相關部份作一個粗略的描述。

考慮一具有n個自由度（degrees of freedom）的力學系統，例如在3維空間的一個封閉容器裏有k個粒子（如：氣體的分子），即$n=3k$的情形，倘若我們已經完全掌握各粒子的質量以及它們間的作用力，則這個系統的瞬時狀態只取決於代表它們位置的n個坐標和所對應的速度。我們也不是一定要取這2n個坐標，譬如，由於某些原因，位置和動量的坐標要遠比位置和速度的來得方便。

在上述的觀點之下，這個系統的一個狀態可以看成在2n維歐氏空間，所謂的相空間（phase space），裏的一個點。隨著時間的過去，力學系統的狀態會跟從適合的物理定律（微分方程）而改變；因此，這系統的整個歷史，過去，現在和將來會以一條軌線（trajectory）呈現在相空間裏。根據古典力學，只要給定這條軌線上的一點（一個瞬時狀態），則在理論上，整條軌線都已經決定好了。但在實際的應用上，我們幾乎不可能有足夠的資料來作這個決定。統計力學的觀念是Gibbs首先提出的，他不再固執於去決定一個狀態（即：相空間的一點），而熱衷在整體狀態統計性的研究（即：相空間的一個子集）。不再期望知道「在時間爲t時系統的狀態是怎樣？」，我們轉而去問「在時間爲t時系統狀態屬於相空間中其一特定子集的機率如何？」最令我們感到興趣是漸近的問題：「當時間t趨近於無窮時，到底有怎樣的情況（可能）發生？」

我們用相空間的點x_t來表示一個固定系統在時間t時的狀態，當t固定時，T_t爲將x_0送到x_t的那個變換，也就是說：$x_t = T_t(x_0)$。顯然，在物理的意義底下，我們會有：$T_{s+t} = T_s T_t$，因此$\{T_t\}$形成一個單參數變換群（one parameter group of transformations）[1]。（這樣的一個群通常被稱爲一個流態（flow）。）統計力學上有一個基本的結果，即Liouville定理[2]，它告訴我們如果適當的選取刻劃相空間的坐標系統，則相空間裏的流態能夠保持所有的體積（ie., 2n維體積）不變。這等於在說流態是由保測變換所構成的；統計力學的基本問題就變成對一族保測變換之漸近性質的研究。

上述的討論中，將一個原來具體的，三度空間的物理架構，生成一頗抽

象，高維度的，具有重要性質的數學模式，（就是流態的保測性）。對於前面的這種過程，我們還有很多非常具體的模型。譬如，考慮一個盛有冰塊與杜松子酒的鷄尾酒器皿，現在加進數滴苦艾酒，用一根棒子很均勻地在攪拌，這樣所造成的力學系統與流態，蠻有趣的，却也適用於解釋我們開始所提到的幾個疑慮。(3)

正如前面我所描述的一樣，遍歷理論是物理問題的數學延展。這門學問有很多很有趣，却並不簡單的定理，並跟數學的其他分支頗有接觸（如機率，拓樸群及Hilbert空間）(4)。與此同時，它也有感傷的一面：將本質上基本的架構弄得晦澀不明。爲了强調它的定理與例子，而減弱它的病態與不合的反例，我決定不作極度推廣的努力。這並不是有意去掩飾它的困難之處，而是去避免那些令我們不太喜悅的場面，畢竟，對一個入門者而言，這是絕對沒有必要的。我會很精確的敍述一些定理，並以正確的方法證明它們，但我也會利用它們來簡化開始看起來頗爲嚴苛的假設。（其中一種方法是在所討論的範圍裏面，舉出一些清楚而尚未解決的問題，却是頗有深度與趣味的問題。）因此，倘若假設某一個測度是有限（或σ－有限）的，或者假設某一拓樸空間具有一可數基以後，對整體有幫助的話，我會毫不猶豫的去做它。

第一個化簡的假設，是將連續的動作演變成離散的程序，我們一直會堅持這個結果。流態的每一個特定的成員，即$T=T_{t_0}$，都是保測變換，（對所有的t_0），由T_t那個群的性質得到$T_{nt_0}=T^n$，n爲任意的整數，正的，負的或零。（$T^0=T_0$爲單位變換。）我們有足夠的理由去假設T_t的漸近性質必須跟T^n的一樣(6)，因此將會把注意力集中在後面那個離散的情況，這是我們的一種偏愛。它其實有些物理的意念；告訴我們如果想對關於流態的漸近性質作一些探討，只須要在一組離散的，等長的時間區間上觀察就可以了。而在數學上的解釋則是將重點放在一些比較基本的觀念上。我們首先會討論的對象是個別的保測變換；而它們所形成的變換群就須要稍待片刻。

在我對這個理論作系統性的討論之前，須要再講一句話，在現代的遍歷

理論中，很多人相信其中一、兩個著名的極限定理是最有意義的細目。我却不以爲然，我們須要學習很多代數性與拓樸性的結果；只有經過拓樸—代數的一般處理以後，複雜的解析事實才會呈現眞正的內涵。

例子

基本的概念是一個測度空間，也就是一集合X，附以由X的一些子集所組成的一個特定的σ－代數，以及定義在這個代數上的測度。不要忘記所謂σ－代數是一對餘集和可數聯集運算封閉的集合族，而非負（可能無限）「可數可加性」集合函數（Countably additive set function）稱爲一個測度。我們所考慮的測度空間都是σ－有限的，這等於說，假設X爲可數個有限測度子集的聯集。主要原因是避免牽涉到Fubini定理和Radon-Nikodym[7]定理時所可能產生的阻滯；在現在σ－有限條件之下，這些定理就可以很圓滑地運用了。

下面我們就給出這種測度空間的一些典型例子：

(1)一個有限維的歐氏空間，附以Borel可測性及Lebesgue測度。

(2)單位區間，跟上面一樣的可測性及測度。

(3)所有序列$x=\{x_n\}$所構成的集合，其中$x_n=0$或1，而n跑完所有的整數。$\{x \mid x_n=1\}$這種類型的集合生成一個σ—代數，可測集是它的元素，而k個不同的這類型集合的交集，賦與測度$\frac{1}{2^k}$，這個條件的確能延拓成一測度。[8]

(4)一具可數基的局部緊緻拓樸群（locally compact topological group）[9]，附以Borel可測性[10]與Haar測度[11]。

可測變換是測度空間之間的變換，而且滿足：可測集的逆像集（inverse image）也一定是可測集。而對一個從X到Y的可測變換T，如果存在一個從Y到X的可測變換S，使ST和TS分別都是單位變換（在它們的定義域上），則稱T是可逆。這個S是由T所唯一決定的；稱爲T的逆變換，記爲T^{-1}。

我們所考慮的可測變換大部份是保測變換，也就是可測集的測度和它逆像集的測度一樣。其實，坦白說，我們感興趣的對象不是眞正的保測變換，

而是這些變換的等價類（equivalence classes）；兩個變換如果最多只有在一個零測度集上相異，則爲等價。爲了使這種等價類的方法通行無阻，我們通常會搬出一個專門術語「認同」（identify）；我打算把兩個保測變換認同若且唯若它們幾乎處處相等。有沒有覺察到如果一個保測變換是可逆的，則它的逆變換也是保測的。而遍歷理論中所討論的變換絕大部份是一個測度空間到它本身的可逆保測變換。

定義在實數線上可測但並不保測的變換的一個典型的例子是：$T(x)=2x$；很容易證明，對每一個Borel集E，都有$m(T^{-1}(E))=\frac{1}{2}m(E)$（當然，m是所討論的測度，現在是Lebesgue測度），一個跟這例子有密切關係而定義在單位區間上的變換是：$T(x)=2x\ (\mathrm{mod}\ 1)$[12]。爲了弄到完全明白，我們考慮半開單位區間〔0，1），而且：

$$T(x)=2x,\qquad \text{當 } 0\leq x<\frac{1}{2}$$

$$T(x)=2x-1,\ \text{當}\ \frac{1}{2}\leq x<1$$

如果$E=(\frac{2}{8},\frac{5}{8})$，則$T^{-1}(E)$是$(\frac{2}{16},\frac{5}{16})$和$(\frac{1}{2}(\frac{2}{8}+1),\ \frac{1}{2}(\frac{5}{8}+1))$的聯集，因此 $m(T^{-1}(E))=\frac{3}{16}+\frac{3}{16}=\frac{3}{8}=m(E)$。同理我們可以證明對所有端點爲二進位有理數的半開區間E，都有$m(T^{-1}(E))=m(E)$，由此容易得到T爲保測變換。T不是1對1的（事實上，它到處都是2對1的），即使去掉一個零測度集也不能使它1對1，因此T是一個不可逆的保測變換。這個變換有下面一個同構（isomorphic，我們尚未定義，心照不宣就好了）的表現形式：對所有絕對值爲1的複數，用Borel測度，且運用就範法（normalized），使一段弧的測度爲其長度的$\frac{1}{2\pi}$倍；然後定義$T(z)=z^2$。[13]

實數線上可逆保測變換一個簡單例子是$T(x)=x+1$。一般地，在一

有限維歐氏空間，令 c 爲一任意向量，並定義T爲T(x) = x + c；更一般地，在一局部緊緻群中，用一左一不變（left-invariant）Haar 測度，令 c 爲這群中任一元素，定義T爲T(x) = cx 。這種一般化的一個很有用的特例是考慮單位圓構成的群得來的；這是一個旋轉 arg c 的變換。它也有一個同構的表現形式，取一介於0與1間的數 c ，在單位區間上定義T，T(x) = x + c (mod 1)，說清楚一點，就是：

$$T(x) = x + c \qquad \text{當} 0 \leq x < 1 - c$$

$$T(x) = x + c - 1 \text{，當} 1 - c \leq x < 1 \qquad (14)$$

我們舉另一堆例子，在一個2維歐氏空間上，定義變換T(x, y) = $(2x, \frac{1}{2}y)$。單位正方形的逆像集爲一底長$\frac{1}{2}$，高2的長方形。其實，同理，每一個長方形的逆像集都是一個有相同面積的長方形，因此T是保測變換；很明顯T是可逆的。讓我們來看看這例子的一般形式，考慮在一有限維歐氏空間上的線性變換T，T的值域是一個子空間，其逆像集爲整個空間，可是眞子空間的測度爲0，因此T爲保測的必要條件爲它是非奇異的（non-singular）。反之，若T爲非奇異，行列式值爲 d，則大家都知道，對每個Borel 集 E，$m(T^{-1}(E)) = m(E) / |d|$，（這個著名的結果很少看到它的證明。讀者可用牽涉 Jacobians 的分析技巧來證得；也有直接的證法，可看Carathe odory，Vorlesungen ueber reelle Funktionen，1927，p.346）。

有限維實向量空間上非奇異線性變換可用其加法向量群的連續自同構（automorphisms）來刻劃。這建議我們考慮在一具左一不變Haar 測度的局部緊緻群上，有一連續自同構T。到底T是否保測呢？我們須比較m(E)和$n(E) = m(T^{-1}(E))$，集合函數n顯然是一個測度；很自然地問，它是個左一不變測度嗎？也就是說：$m(T^{-1}(xE))$等於$m(T^{-1}(E))$嗎？答案是肯定的，因爲$T^{-1}(xE)$是$T^{-1}(E)$的一個左平移；事實上，經過簡單的計算，得到$T^{-1}(xE) = (T^{-1}(x^{-1}))^{-1}T^{-1}(E)$。(15) 由Haar 測度的唯一性，得證$m(T^{-1}(E))$爲m(E) 的一個常數倍。一般來

講，這是我們能夠做到最好的了；非奇異線性變換告訴我們自同構不一定是保測的。假如，眞的那麼僥，X是個緊緻群，則m(X)是有限，因此令E爲X就可以計算出那個比例常數；但$T^{-1}(X)=X$，故此常數爲1，所以T是保測的。

這種群有一個很有趣的特例——環面（torus），也就是兩個單位圓的Cartesian 乘積。具體地來說，這個群的元素是由兩個模爲1的複數構成的偶對（u，v）；而其運算是逐標（coordinatewise）乘法[16]。容易證明這個群裏一般的連續自同構可以用一個二階的「單位模」（unimodular）方陣來刻劃，即一個頂點元素爲整數，行列式值爲±1的方陣；若 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 爲此種方陣，則它對應的自同構T定爲：

$$T(u,v)=(u^a v^b, u^c v^d) \qquad (17)$$

令X爲前面所提過包含所有數列$x=\{x_n\}$，$n=0,\pm 1,\pm 2,\cdots\cdots$，的空間；令T爲對足碼單位移換（unit shift）所誘導出的那個變換，也就是：$T(x)=y=\{y_n\}$，其中$y_n=x_{n+1}$。這個變換是保測及可逆的。如果x爲單向數列空間（unilateral sequence space），即x的元素爲數列$\{x_n\}$，$n=0,1,2,\cdots\cdots$，跟上面T的同一方程式定義出一個保測但不可逆（2對1）的變換。

單向數列空間與單位區間之間有一個簡單的變換S；S將由0與1構成的數列$\{x_n\}$映到二進位表示法爲 $0.x_1x_2x_3\cdots\cdots$的那個數。S是保測而且可以看成1對1的。嚴格說來S並不是1對1的，原因是有些二進位有理數有兩種小數展開形式。那些會被S映到二進位有理數的數列所形成的集合，和二進位有理數的集合有相同的基數；都是可數無限的。如果我們適當的去掉這些數列，重新定義S，結果得到一從這數列空間到單位區間的可逆保測變換。這種變換的存在告訴我們在測度理論上，這兩個空間的結構是同構的。而同構變換（就是S）將單向移換T對應到單位區間上的一個保測變換T′；T′可以定義成：$T'=STS^{-1}$。由S和T的定義說明T′其實是我們的老朋友：

$$T'(x) = 2x \quad (\text{mod } 1), \text{ a.e.} \qquad (18)$$

另一方面，雙向數列空間與單向數列空間跟其本身的Cartesian乘積間有一個很自然的對應；也就是：

將 $\{\cdots\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots\cdots\}$

映到 $(\{x_0, x_1, x_2, \cdots\cdots\}, \{x_{-1}, x_{-2}, \cdots\cdots\})$

很容易看出這個對應是可逆的保測變換，因此是一個測度理論上的同構，記爲P。若Q爲S與它本身的Cartesian乘積（所以$Q(x, y) = (S(x), S(y)$，其中x，y爲單向數列），則合成變換QP爲雙向數列空間到單位正方形的同構。這個同構將雙向移換對應到正方形上的一個可逆保測變換T″。再次檢查定義知道T″跟我們的老朋友也密切相關；它就是：

$$T''(x,y) = \left(2x, \frac{1}{2}y\right) \qquad \text{當} \ 0 \leq x < \frac{1}{2}$$

$$T''(x,y) = \left(2x, \frac{1}{2}(y+1)\right) \qquad \text{當} \ \frac{1}{2} \leq x < 1 \qquad (19)$$

（這方程當然要在（mod 1）下，才幾乎處處成立。）

T′也可以用幾何方法描述如下：先將單位正方形上點（x，y）映到$(2x, \frac{1}{2}y)$，得到一個底邊爲〔0，2），左邊爲〔0，$\frac{1}{2}$）的長方形；將這長方形的右半邊切開（底邊爲〔1，2）的），再用平移把它搬至左半邊的上面。這不是很像在搓麵團嗎？T″就叫做「麵包師變換」（Baker's transformation）。(20)

遍歷性

如果T爲X上的保測變換，而X能夠表示成兩個互不相交，在T作用下不變的正測度集E，F的聯集，則探討T在X上的性質可以分別從E和F著手。在這種情況下，T稱爲可分解的（decomposable），所以眞正富有意義的是那些不可分解的變換——通常稱爲度量可遷的（metrically transitive）或是遍歷的（ergodic）。一個變換可以看成它所作用的空間上的一次遷涉，而遍歷性爲應「妥當遷涉過程」這個自然的須求而定的其中一

種恰當形式。

爲了給出一些遍歷變換的例子，先列舉遍歷性的另外一些定義形式可能對我們比較方便。第一種形式是極爲明顯的：T是遍歷的充要條件是它只有一些無聊的不變集。也就是等價於：

「E是可測，T－不變集

$\Rightarrow m(E)=0$ 或 $m(X-E)=0$」

（不要忘記不變性的定義，「E爲T－不變集 $\Rightarrow T^{-1}(E)=E$ 」；這表示 $x\in E$ 若且唯若 $T(x)\in E$ 。一個函數f是T－不變的意思是「$f(T(x))=f(x)$，$\forall x$ 」。顯然E是T－不變的充要條件是它的特徵函數是T－不變的，一般來講，我們所謂的「不變」其實的意思是「幾乎處處不變」，因此，譬如函數的不變性就是「$f(T(x))=f(x)$幾乎處處成立。」）

遍歷性一個有用的等價形式是：

「T是遍歷 $\Rightarrow$每個可測，T－不變函數都是常數函數」。 其中一個方向 的證明是顯然的：如果沒有非常數的T－不變函數，當然就找不到非無聊性 的T－不變子集。（考慮特徵函數。）反之，假設T是遍歷的；須要證明倘 若f是可測T－不變函數，則f是常數函數。

定義 $X(k,n)=\{x\mid \frac{k}{2^n}\le f(x)<\frac{k+1}{2^n}\}$，

f的T－不變性隱涵 $X(k,n)$ 是T－不變集。而T的遍歷性就表示：「對固定的n，除了一個k以外，$m(X(k,n))=0$」。對所有的n，挑那個比較大的 $X(k,n)$ 作交集，我們所須要的結果就馬上得到了。[21]

對 $m(X)<\infty$ 的情形來說，我們有：

「T是遍歷 $\Leftrightarrow$每一個在L_1中T－不變函數都是常數函數」（或者，像我們，可以只取L_2的函數）；

這是因爲每個可測集的特徵函數都是可積分的。

在整數空間上的平移T，$T(x)=x+1$ ，是遍歷的；但另一個平移T，$T(x)=x+2$，就不是了。（偶數集是T－不變的）而 $T(x)=x+1$ 這個平移在實數線上也不是遍歷的；這是因爲

$\bigcup_{n=-\infty}^{\infty}(n, n+\frac{1}{2})$ 是一個T－不變集（非無聊性的）。

若X爲單位圓群（即：包含所有絕對值爲1的複數所成的集合），而 $c \in X$，　定義：$T: X \to X$，　「$T(x) = cx$」，

則T的遍歷性是跟c有關的。

如果c爲1的一個根，即存在正整數n，使$c^n = 1$，則T不是遍歷的；事實上，若$f(x) = x^n$，則f是一個非常數、可測的T－不變函數。

如果c不是1的一個根，則T是遍歷的。其中一種證法是這樣的：

定義　$f_n(x) = x^n$，$n = 0, \pm 1, \pm 2, \cdots\cdots$ f_n 能夠扶正成L_2的一個完備的就範直交集（complete orthonormal set），因此若$f \in L_2$，則$f = \sum_n a_n f_n$，其中這個級數爲二次平均收斂，再定義算子U：

$$U(f(x)) = f(T(x));$$

而　$U(f_n) = c^n f_n$，　故　$U(f) = \sum_n a_n c^n f_n$。

若f是T－不變的，則　$\forall_n$，　$a_n = a_n c^n$

因此：只要$n \neq 0$，就有　$a_n = 0$

這表示在L_2中的每個T－不變函數都是常數函數，所以T是遍歷的。

更一般地，如果X爲一個具有可數基的緊緻交換群，而

$T(x) = cx$，　其中　$c \in X$，則：

「T是遍歷 $\Leftrightarrow$ 點列$\{ c^n \mid n = 0, \pm 1, \pm 2, \cdots\cdots \}$在X中稠密」

這件事情的證明是離開我們的主題一個有趣的插敍；大致如下：

引理　如果測度空間X是一個具可數基的拓樸空間，而且每個非空開集都有正測度

若T是遍歷的保測變換，

則對X中幾乎所有的點x，

x的軌跡$\{ T^n(x) \}$在X中稠密。

證明：　否則，存在一基元素、非空開集G，滿足：

$$「x \in \bigcap_{n=-\infty}^{\infty} (X - T^n(G))」。$$

上面那個交集是跟G不相交的一個T一不變集，

而 $m(G) > 0$，

所以
$$m\left(\bigcap_{n=-\infty}^{\infty} (X - T^n(G))\right) = 0$$

如果 $x \in \cup \{ \bigcap_{n=-\infty}^{\infty} (X - T^n(G)) \mid G$ 爲一基元素的開集 $\}$

則x就有一個稠密的軌跡。（上面的集合零測度）

在引理中，稠密性是遍歷性的一個必要條件，但它却並不是充分的。讓我們來舉一個反例吧：

先取 $T : 〔0, 2) \to 〔0, 2)$

其中要求 $〔0, 1)$，$〔1, 2)$ 均不是T一不變集，且當T分別限制在它們上面時是遍歷的。我們的例子是這樣的，首先在 $〔0, 2)$ 上定義一個拓樸（當然有別於「常用的拓樸」（usual topology）），使得子區間 $〔0, 1)$ 及 $〔1, 2)$ 均在 $〔0, 2)$ 中稠密[22]，且所得出的拓樸、測度空間滿足引理中的條件。爲此，考慮平面上的一正方形，在其中挑出兩組相異的半開區間，每組都只有可數個，都在正方形中稠密，它們的聯集跟 $〔0, 1)$ 與 $〔1, 2)$ 有一個很自然的1-1對應，這個對應就在 $〔0, 2)$ 中誘導出一個拓樸。要不然的話，透過這個對應，我們也可以定義正方形上的一個測度，給予那些不相干點的集合的測度爲零，再定義一個變換，使得：

對不相干的點 $x \to x$

對其他的點 $x \to T(x)$。

現在假設T爲一旋轉（即，$T(x) = cx$，在一個具有可數基的緊緻交換群上）。如果T是遍歷的，則由引理最少有一點 x_0 的軌跡是稠密的。我們也知道變換 $x \longrightarrow x x_0^{-1}$ 是同胚，故它將 x_0 的軌跡 $\{ c^n x_0 \}$ 變成一個稠密序列；它就是c的冪方 $\{ c^n \}$。

反之，假設 $\{ c^n \}$ 爲稠密，若f爲X的一個「特徵」[23]（charac-

ter），即一個到單位圓群的連續同態（homomorphism），

則　$f(cx)=f(c)f(x)$，

因此，f 是由T所誘導出來的那個酉算子的一個固有向量（proper vector），而特徵構成L_2的一個完備就範直交集，故L_2中的T－不變函數都可以根據它們來展開；但是已知酉算子對應於不同固有值的固有向量互相垂直，所以L_2中每個T－不變函數（即具有固有值1的固有向量）能夠表成固有值爲1的特徵的線性組合。剩下來只須要證明唯有「主特徵（principle character）」[24] 的固有值是1。事實上，若f爲一特徵，且

$f(cx)=f(x)$　幾乎處處成立，

由連續性，　$f(cx)=f(x)$　到處都成立，

因此　$f(c^n x)=f(x)$，$\forall x$，$\forall n$

令$x=1$，我們就得到所要的結果。

再來看一下定義在環面上的旋轉T，其中

$T(x,y)=(bx,cy)$

我們可以利用上述的拓樸技巧或者是Fourier展開的方法證得：

T是遍歷

⇔乘子（multipler）b，c的坐標是「數性」獨立的（integrally independent；意思是：

$\forall$整數m，n，　$b^n c^m=1 \Rightarrow n=m=0$）

定義在有限維歐氏空間上，具有行列式值1的線性變換能不能是遍歷的呢？答案是否定的。其中一種證法是將線性變換的固有值理論應用到複向量空間上；首先要把所討論的空間X「複性化」（complexifed），也就是考慮Cartesian乘積$X\times X$；定義逐標的向量加法，以及「複純量乘積」：

$(a+ib)(x,y)=(ax-by,bx+ay)$。

結果我們得到一個複維度爲n的複向量空間$\widetilde{X}$（n爲X的實維度），當然$\widetilde{X}$的實維度是2n；而這個2n維的實向量空間包含n維子空間X。

定義　$\widetilde{T}:\widetilde{X}\to\widetilde{X}$

$\widetilde{T}(x,y)=(T(x),T(y))$，

則$\widetilde{T}$是一個具行列式值1的複線性變換。

令$c_1, c_2, \cdots\cdots, c_k$爲$\widetilde{T}$ 的固有值，

分別具重覆度（multiplicities）$n_1, n_2, \cdots\cdots, n_k$，(25)

再令$z_1, z_2, \cdots\cdots, z_k$爲所對應$\widetilde{T}$ 的非零固有向量。

注意，由$\widetilde{T}$ 的定義，向量z_j是$\widetilde{X}$上的複線性泛函（functionals），而$X \subset \widetilde{X}$，所以z_j可以定義在X上；

$$令 \quad f(x) = \prod_{j=1}^{k} (z_j(x))^{n_j} \ ,$$

則f是在T 作用下不變的，（不要忘記所有固有值連同它們重覆度的乘積就是T的行列式值。）(26)可是f却不是一個常數函數。事實上，f僅在z_j的核空間（null-spaces）的聯集上爲零，而z_j的實部與虛部均是X上的實線性泛涵，所以這些核空間的聯集只是有限個維度低於n的子空間的聯集，但f是連續的，由上述的討論得f不是幾乎處處爲常數的。（這裏所用到的「複性化」技巧，主要是爲了簡化我們的步驟，並得到較多的資料，這是考慮矩陣代替線性變換的結果。）

上面的事實能作如何的推廣呢？在一個局部緊緻而非緊緻的群上的自同構會否是遍歷的保測變換呢？我們並不知道結果；只有在緊緻的情形已經做了一點東西。稍後我們就會考慮緊緻群上的自同構。

下面一個例子我將考慮T爲適當的一個數列空間上的單向或雙向移換；無論那一種情形，T都是遍歷的。假設E是一個可測的T－不變集，因爲這裏的測度決定於一些相差有限個坐標的集合上，所以存在一個「有限維」的集合A，使A逼近E，意思是對稱差（或Boolean和）$E \triangle A$的測度很小(27)，特別，它表示$m(E)$接近$m(A)$。但A是由一個含有限個坐標的集合決定的，故當n充分大時，集合$B = T^{-n}(A)$ 是由另一個不相交的坐標集合決定，因此：

$$m(A \cap B) = m(A)m(B) \ 。$$

而T和它的冪方都是保測變換，E是不變集，由於$m(E \triangle A)$很小，故$m(E \triangle B)$也是的，所以$m(E \triangle (A \cap B))$是很小的，也就是$m(A)$，

$m(B)$ 和 $m(A)m(B)$ 都很接近 $m(E)$；換句話說，$m(E)$ 很接它自己的平方，而這種接近的程度是任意的，因此 $m(E)=(m(E))^2$，得證。由此證明「雙倍」變換（$T(x)=2x \pmod 1$）和麵包師傅的那個變換均是遍歷的。

除了整數空間的單位平移以外，我們上面所談的遍歷變換都是作用在有限測度空間上的。要在一個無限測度的「非實心」（non-atomic）空間上造一個遍歷變換並不是一件輕鬆的事情。（非實心的意思每個正測度集都包含一個正測度但測度較小的子集。）[28] 我將會在實數線上造一個這種變換。其實那個空間最方便的表現形式並非實數線而是平面上的一組線段；很明顯的，這些線段會連起來成爲實數線，下面的方法對建構例子有很廣泛的用途。

令 $\{a_n\}$ 爲一遞減的正數列，$a_0=1$，

T_0 爲單位半開區間上的可逆遍歷變換，

X_n 爲長 a_n 的半開區間，置在平面上與水平軸平行；

$X=\cup X_n$，[29]

若 $\sum a_n$ 發散，則顯然 X 在測度上與半線或整條數線同構。現在定義：

$T: X \to X$

$T(x,y)=(x,y+1)$，如果有意義的話

$T(x,y)=(T_0(x),0)$，不然（$y=n, x \geq a_{n+1}$）

顯然 T 是 X 上的可逆保測變換，設 E 爲一在 T 作用下的不變集，令 E_0 爲 E 和「底區間」（base interval，水平軸上的單位區間）的交集，則 E_0 是在 T_0 作用之下不變的，故 E_0 或它在底區間的餘集測度爲零；所以 T 是遍歷的。[30]

〔譯註〕

註〔1〕：$(T_t)^{-1} = T_{-t}$ ，$T_t T_0 = T_0 T_t = T_t$

通常，視個別情況，還要求：

$$T_t(x) \xrightarrow{t \to 0} T_0(x) = x \ , \ \forall x$$

或者 $(t, x) \longrightarrow T_t(x)$ 可微分（最少對 t）

註〔2〕：有興趣的可參考統計力學的書，或者：

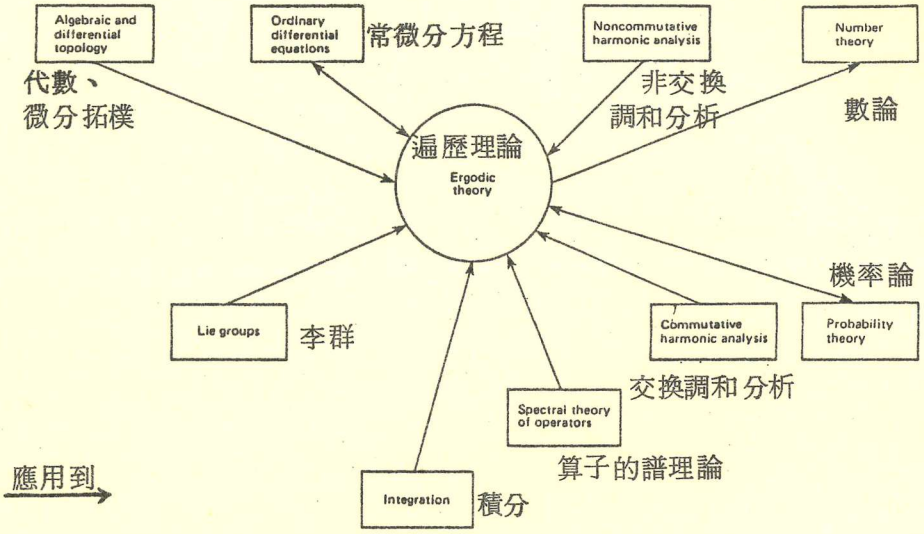
Colin J. Thompson：

Mathematical Statistical Mechanics，

Princeton University press pp.16-18，211-213.

註〔3〕：

註〔4〕：法國數學家J. Dieudonne 在其A Panorama of pure Mathematics有以下的指示：



註〔5〕：$T = T_{t_0}$ ；

$$T_{nt_0} = T_{t_0+t_0+\cdots\cdots+t_0} \quad (\text{n個 } t_0)$$
$$= T_{t_0} T_{t_0} T_{t_0} \cdots\cdots T_{t_0}$$
$$= T\,T\,T \cdots\cdots T = T^n$$

註〔6〕：譬如，若 $t \longrightarrow T_t(x)$ 連續

則 $\lim_{t\to\infty} T_t(x) = \lim_{n\to\infty} T^n(x)$ ，（存在的話）

註〔7〕：參考實變數函數論的書，如：

H.L. Royden, Real Analysis pp.269-270, 238-240.

註〔8〕：考慮機率空間 $(X_n, 2^{X_n}, \mu_n)$ ，$n = 0, \pm 1, \pm 2, \cdots\cdots$

其中 $X_n = \{0, 1\}$ ， $\mu_n(\{0\}) = \mu_n(\{1\}) = \frac{1}{2}$

造 $X=\prod_{n=-\infty}^{\infty} X_n$ ，

在其中用乘積測度（product measure）：

$$m\left(\bigcap_{i=1}^{\ell}\{x=\{x_n\}\mid x_{k_i}=1\}\right),\ k_1<k_2<\cdots\cdots<k_\ell$$

$$=m(\cdots\cdots\times X_{k_1-1}\times\{1\}\times X_{k_1+1}\times\cdots\cdots\cdots\times X_{k_\ell-1}\times\{1\}\times X_{k_\ell+1}\times\cdots\cdots)$$

$$=\mu_{k_1}(\{1\})\cdot\mu_{k_2}(\{1\})\cdots\cdots\mu_{k_\ell}(\{1\})$$

$$=\frac{1}{2}\cdot\frac{1}{2}\cdots\cdots\frac{1}{2}=\frac{1}{2^\ell}$$

註〔9〕：Y是拓樸群的意思是：

Y是一個Hausdorff 拓樸空間，也是一個群，而且，運算 $(a,b)\longrightarrow ab^{-1}$ 是連續的。

註〔10〕：拓樸空間的Borel 集就是由開集所生成的 σ－代數的元素。

註〔11〕：m爲拓樸群Y上的一個左－Haar 測度的意思是：

(i) m是一個測度

(ii) $m\neq 0$

(iii) 每個緊緻集C 都可測，且 $m(C)<\infty$

(iv) 若E 可測，則 $\forall x$ ，xE 可測

且 $m(xE)=m(E)$

局部緊緻拓樸群上都有左－Haar 測度，每兩個左－Haar 測度 m_1 ，m_2 都滿足：

$\exists k>0$ $\quad m_1=km_2$

（若爲機率空間，則此 $k=1$ ，左－Haar 測度唯一存在）。

註〔12〕：（mod 1）的意思是取（正）小數部份，

所以 $\quad T(x)=2x\ (\mathrm{mod}\ 1)$

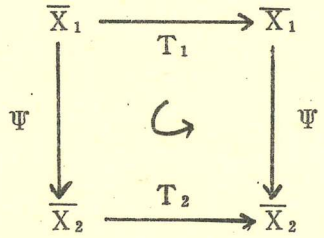
可以寫作 $\quad T(x)=2x-[2x]$

註〔13〕：$T_1:X_1\to X_1$ ，$T_2:X_2\to X_2$ 爲同構的充要條件是：

$\exists\ \overline{X}_1 \subset X_1$, $\overline{X}_2 \subset X_2$, $T_1^{-1}(\overline{X}_1) = \overline{X}_1$, $T_2^{-1}(\overline{X}_2) = \overline{X}_2$

$\exists$ 可逆保測變換 $\Psi : \overline{X}_1 \to \overline{X}_2$

使得 $\Psi T_1 = T_2 \Psi$



這裏，$X_1 = \overline{X}_1 = [0, 1)$, $X_2 = \overline{X}_2 = \{ z \mid |z| = 1 \}$

$T_1(x) = 2x \pmod 1$, $T_2(z) = z^2$

$\Psi(x) = e^{2\pi i x}$,

註〔14〕：如上面〔12〕的，這裏同樣 $\Psi(x) = e^{2\pi i x}$

而後面那個 c 其實是 $\frac{1}{2\pi} \arg c$

註〔15〕： $y \in T^{-1}(xE)$

$\Rightarrow T(y) \in xE$

$\Rightarrow x^{-1} T(y) \in E$

$\Rightarrow T(T^{-1}(x^{-1}) y) = T(T^{-1}(x)) T(y) \in E$

$\Rightarrow y \in (T^{-1}(x^{-1}))^{-1} T^{-1}(E)$

註〔16〕：$(a, b)(c, d) = (ac, bd)$

註〔17〕：令 T 爲環面上的一個連續自同構，

ω 爲一複數，$\omega^k = 1$, $\omega \neq 1$, k 爲正整數。

則 $(T(1, \omega))^k = T(1, \omega^k) = T(1, 1) = (1, 1)$

所以 $T(1, \omega) = (\omega^b, \omega^d)$, b , d 爲整數

若 θ 是另一個複數，$\theta^\ell = 1$, $T(1, \theta) = (\theta^x, \theta^y)$

則 $T(1, \omega\theta) = T(1, \omega) T(1, \theta)$

$= (\omega^b, \omega^d)(\theta^x, \theta^y)$

$= (\omega^b \theta^x, \omega^d \theta^y)$

但 $\omega\theta$ 也是1的根，從上面可推得：

$\exists b, d$ $\forall$ 1的根ω，$T(1,\omega)=(\omega^b, \omega^d)$

但所有1的根的這個集合在單位圓上稠密，

因此，若 $|z|=1$，則 $T(1,z)=(z^b, z^d)$

同理，$\exists a, c$， $\forall |z|=1$，$T(z,1)=(z^a, z^c)$

故 $T(u,v)=T(u,1)T(1,v)=(u^a v^b, u^c v^d)$

又 T^{-1} 也是圓環上的自同構，

故 $T^{-1}(x,y)=(x^e y^f, x^g y^h)$

而 $(u,v)=T^{-1}(T(u,v))$

$=T^{-1}(u^a v^b, u^c v^d)$

$=C(u^a v^b)^e (u^c v^d)^f, (u^a v^b)^g (u^c v^d)^h)$

$\therefore$ $ae+cf=1$，$be+df=0$

$ag+ch=0$，$bg+dh=1$

也就是 $\begin{pmatrix} e & f \\ g & h \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

而a，b，c，d，e，f，g，h都是整數

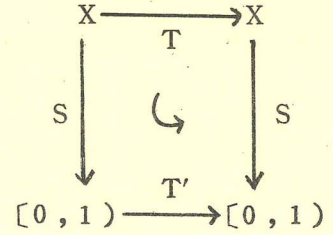
$\therefore$ $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\pm 1$

還有其他的證法，請參考：

Peter Walter An Introduction to Ergodic Theory

(G.T.M. 79) pp.14-16.

註〔18〕：令X爲單向數列空間，（重新定義後）



$;T'=STS^{-1}$

若 $x\in$〔0，1）具有二進位表示法 $0.x_1 x_2 \cdots\cdots$

則 $T'(x) = STS^{-1}(0.x_1x_2\cdots\cdots)$

$= ST\{x_1, x_2, \cdots\cdots\}$

$= S\{x_2, x_3, x_4, \cdots\cdots\}$

$= 0.x_2x_3\cdots\cdots$

若 $x_1 = 0$， 即 $x \in [0, \frac{1}{2})$

則 $0.x_2x_3\cdots\cdots = 0.0x_2x_3\cdots\cdots$的兩倍

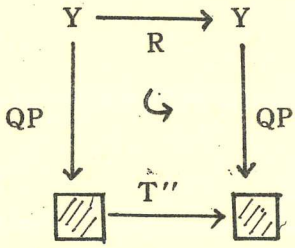
$= 2x$

若 $x_1 = 1$， 即 $x \in [\frac{1}{2}, 1)$

則 $0.x_2x_3\cdots\cdots + 1 = 0.1x_2x_3\cdots\cdots$的兩倍

所以 $T'(x) = 2x \pmod 1$

註〔19〕：令Y爲雙向數列空間，R爲其上的移換



$; T'' = (QP)R(QP)^{-1}$

$T^*(0.x_0x_1x_2\cdots\cdots, 0.x_{-1}x_{-2}\cdots\cdots)$

$= (QP)R(\{\cdots\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots\cdots\})$

$= (0.x_1x_2\cdots\cdots, 0.x_0x_{-1}x_{-2}\cdots\cdots)$

若 $x_0 = 0$

$0.x_1x_2\cdots\cdots = 0.0x_1x_2\cdots\cdots$的兩倍

$0.0x_{-1}x_{-2}\cdots\cdots = 0.x_{-1}x_{-2}\cdots\cdots$的一半

若 $x_0 = 1$

$0.x_1x_2\cdots\cdots + 1 = 0.1x_1x_2\cdots\cdots$的兩倍

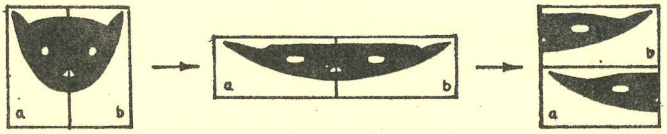
$0.1x_{-1}x_{-2}\cdots\cdots$的兩倍$= 0.x_{-1}x_{-2}\cdots\cdots + 1$

所以 $T''(x,y)=(2x,\frac{1}{2}y)$，$0\le x<\frac{1}{2}$

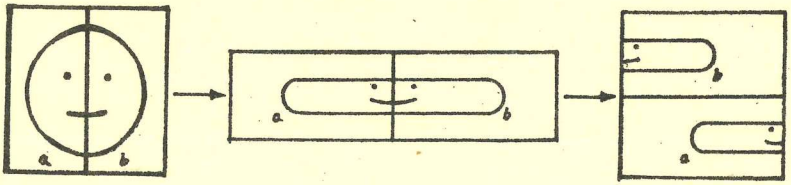
$T''(x,y)=(2x,\frac{1}{2}(y+1))$，$\frac{1}{2}\le x<1$

註〔20〕：Baker's transformation：

Arnold型的：



Ornstein型的：



註〔21〕：$\forall n$，$\exists k_n \ni X(k_n,n)=f^{-1}([\frac{k_n}{2^n},\frac{k_n+1}{2^n}))$滿足

$$m(X)=m(X(k_n,n))$$

作交集 $\bigcap_{n=1}^{\infty} X(k_n,n)$

則 $m(\bigcap_{n=1}^{\infty} X(k_n,n))=m(X)$

但 $\bigcap_{n=1}^{\infty} X(k_n,n)=\bigcap_{n=1}^{\infty} f^{-1}([\frac{k_n}{2^n},\frac{k_n+1}{2^n}))$

$$=f^{-1}(\bigcap_{n=1}^{\infty}[\frac{k_n}{2^n},\frac{k_n+1}{2^n}))$$

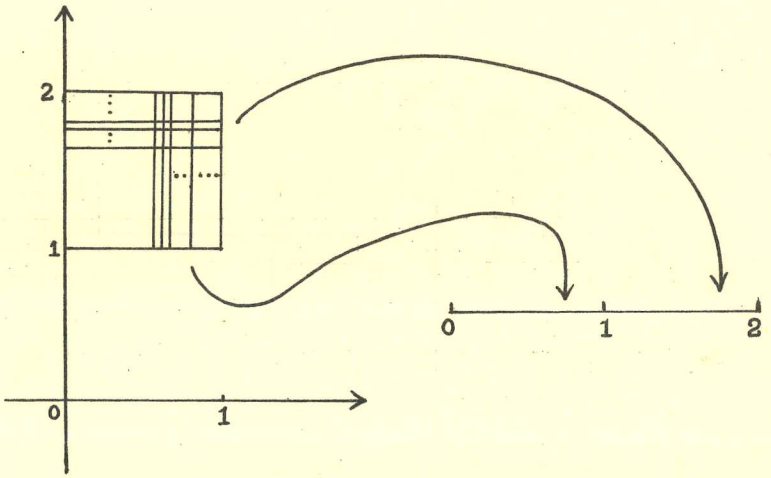
而 $\left[\frac{k_n}{2^n}, \frac{k_n+1}{2^n}\right)$ 的長度 $\longrightarrow 0$ ，

但其交集爲連通的非空集，故爲單點集 $\{c\}$ ，

即 $f(x)=c$ ，a.e.

註〔22〕：只須在 $[0,2)$ 取拓樸如 $\{\phi, [\frac{1}{2},\frac{3}{2}), [\frac{1}{4},\frac{5}{4}), [0,2)\}$ 就可以了。

Halmos 的意思大概是這樣子的：



$T: [0,2) \to [0,2)$

$T^{-1}[0,1)=[0,1)$, $T^{-1}[1,2)=[1,2)$

$T|_{[0,1)}$, $T|_{[1,2)}$ 是遍歷的

則對幾乎所有的 $x\in[0,1)$ ，

$Cl_{[0,1)}(\{T^n(x)\mid n=0,\pm1,\pm2,\cdots\})=[0,1)$

故 $Cl_{[0,2)}(\{T^n(x)\mid n=0,\pm1,\pm2,\cdots\})$

$=[0,2)$

同理，對幾乎所有的 $x\in[1,2)$ ，事實上對所有的 $x\in[0,2)$ ，x 的軌跡 $\{T^n(x)\}$ 都會在 $[0,2)$ 中稠密。

但顯然T不是遍歷的。

註〔23〕：關於Character，可參考：

Peter Walter Introduction to Ergodic Theory pp.12-13.

註〔24〕：常數值的Character.

註〔25〕：$\widetilde{T}^*:\widetilde{X}^*\to\widetilde{X}^*$的其中一個表現矩陣爲：

$$A=\begin{pmatrix} c_1 & & & & \\ & \ddots & & & \\ & & c_1 & & \\ & & & c_2 & \\ & & & & \ddots \\ & & & & & c_k \\ & & & & & & \ddots \\ & & & & & & & c_k \end{pmatrix}_{n\times n}$$

c_1 重複 n_1 個，c_2 重複 n_2 個，……，c_k 重複 n_k 個；$n_1+n_2+\cdots\cdots+n_k=n$

註〔26〕：

$$f(x)=\prod_{j=1}^{k}(z_j(x))^{n_j}$$

$$f\circ T(x)=\prod_{j=1}^{k}(z_j(T(x)))^{n_j}$$

$$=\prod_{j=1}^{k}(T^*(z_j)(x))^{n_j}$$

$$=\prod_{j=1}^{k}(c_j z_j(x))^{n_j}$$

$$=\prod_{j=1}^{k}c_j{}^{n_j}\cdot\prod_{j=1}^{k}(z_j(x))^{n_j}$$

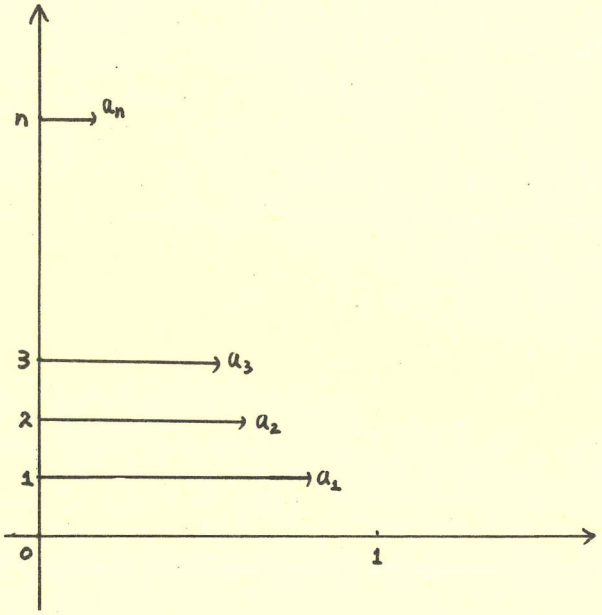
$$=(\det A)f(x)$$

$$=f(x)$$

註〔27〕：E△A在原文中是E＋A

註〔28〕：因此每個單點集的測度均爲零。

註〔29〕：



註〔30〕：由T的定義，

若E_0的測度是零的話，則 $E \cap [1, a_1)$的測度也是零，

同理 $m(E \cap [n, a_n)) = 0, \forall_n$

所以 $m(E) = 0$

若$[0, 1) - E_0$的測度是零，則$[1, a_1) - E$測度也是零，同理 $m([n, a_n) - E) = 0, \forall_n$

所以 $m(X - E) = 0$

後記：

這年頭「讀」純數的人愈來愈少了，大夥兒都去搞一些比較有「價值」的東西，令到風光了好幾千年的數學也有人才難求之苦。哀哉！

有人不信邪，明知重現以往歷史的測度（機率）理論上爲0，但仍堅信

，只要勇往直前，科技有尋根的一天。畢竟，地球是圓的。阿彌陀佛！

最後，但願佛祖「慈悲」，不要讓我研究所的考試全軍盡墨，使我得繼續接受數學的「制裁」。

——1985.4. 遠征清華試刀的前幾天——

3. 亂數(Random Number)產生之探討

指導老師：黃登源
四　甲：古思明

第一章　導論（Introduction）

亂數的產生通常分爲二大類：第一種是用機器，例如：將一個轉輪等分成10份，並標以0～9，然後用一根指針指示，當高速旋轉停止指針所在的位置。第二種方式是使用算術運算的方式於電算機上操作，本文中所要探討的就是第二類。利用算術運算的方法來產生亂數的形式通常是：令 $f: x \to f(x)$，$0 \leq x < m$，x 是整數，$0 \leq f(x) < m$。如果我們將x標上足標，那我們就可以得到形如：$f(x_n) = x_{n+1}$，$0 \leq n < m$，$0 \leq x_{n+1} < m$ 的方式。而若令 $x_i = x_k$，顯然 $f(x_i) = f(x_k) \Rightarrow f(x_{i+1}) = f(x_{k+1})$ ……，故已形成循環現象，而其循環最長週期爲m。

雖然第二類方式所產生的亂數無法項項不相干，（因至少除 x_0 初值，以外，其餘諸數皆是其先數所產生），但所得的數序若能通過統計檢定，則其效果也是相當不錯的。近年來，亂數廣泛的應用在摹擬，隨機樣本、數值分析，決策製定等方面。好的亂數數列從使用上來看，不僅要有好的“統計性質”，也須具備①計算上的效率②不浪費電算機記憶空間③複製同樣的數列具方便性。而由上可知，第二類方式所產生不但能節省記憶空間，而且複製或重複使用一個數序也很方便。

我們先介紹一些名詞：

①在 $f(x_n) = x_{n+1}$ 中的初值 x_0 叫 seed。

②週期：在是列中最小循環數目。如{1，2，5，8，4，1，2，5，8，4 ……}則週期爲5。

③在 $f(x_n) = x_{n+1}$ 中，若週期爲 m，則稱爲全週期（full period）或最大週期。

在全週期發生時，0，1，2，…，m−1 出現的機率通通一樣（當然，是指一週期內若令之爲 $\{x_0, x_1, \cdots\cdots, x_{m-1}\}$，而設 $U_i = \frac{x_i}{m}$，則我們得到一個對應的數列 $\{U_0, U_1, \cdots\cdots, U_{m-1}\}$ 佈於〔0，1）上，當然，我們希望 $\{U_i\}$ 能均勻分佈於〔0，1）上，如此就可擁有 U〔0，1）：uniform distribution on〔0，1）的好性質。爲了達到此目的我們 m 一定要充份大，如此 $\{U_0, U_1, \cdots\cdots, U_{m-1}\}$ 才能稠密（dense）分佈於〔0，1）。

本文先介紹 6 種主要的不同方法，後介紹一種統計檢定的方法，並以亂數應用於數值分析上做一結束。

第二章　各種主要方法介紹

第一節　（Midsquare Method（MID. M））

自從電算機問世不久，科學家們開始嘗試使用它來產生亂數，有許多是效率不高的方法，（例如：將亂數表輸入電算機），而 John Von Neumann 首先於 1946 年左右提出一個相較之下爲高效率的方法——Midsquare Method，其方式如下：假設，我們要產生一 4 位阿拉伯數字的亂數數列，且設初值是 1985，欲得下一位亂數只要將 1985 平方＝ 3940225，然後看看是否爲 8 位數，若不是由其前補上 0，ie　，03940225，取中間 4 位數字＝ 9402，如此步驟重複實行，即可得亂數數列。

x_n	x_n^2	Mid. M	x_{n+1}
1009	01018081		0180
0180	00032400		0324
0324	00104976		1049
1049	01100401		1004
1004	01008016		0080
⋮	⋮		⋮

而Mid. M被證明無法提供足夠多的亂數。實際上，若欲產生具有2n個位數的亂數，而其平方後所取的亂數左邊要n位0，則再經Mid. M，顯然會漸趨於0，例如本例，在x_5時爲0080，則平方後所得是0064，……退化到零。雖然，仍有人找出退化前的75000個亂數（見Knuth p-4），但一般都建議放棄Mid. M。

第二節　線性同餘法（Linear Congruential Method（L.C.M））

D.H. Lehmer在1949年首先提出L.C.M是現階段廣被使用的方法之一，其方式易懂，結構也很容易。其型式如下，令$f(x)=(ax+c)\bmod m$，其中，m是模數（modulus），a是乘數（multiplier），c是增量（increment）。

如果我們給x一個足標，則我們可有$f(x_n)=(ax_n+c)\bmod m=x_{n+1}$ …… ①　　$n=0,1,2,\ldots\ldots$（並設初值爲x_0），我們先看一個例子，以便了解其操作情形。

Ex1，令$a=3$，$c=1$，$m=8$，$x_0=0$

$\Rightarrow x_1=1$，$x_2=4$，$x_3=5$，$x_4=0$，$x_5=1$，……

Ex2，令$a=1$，$c=1$，$m=8$，$x_0=0$

$\Rightarrow x_1=1$，$x_2=2$，$x_3=3$，$x_4=4$，$x_5=5$，$x_6=6$，$x_7=7$，$x_8=0$

Ex1，的週期是4，Ex2，的週期是8。

由$x_{n+1}=(ax_n+c)\bmod m$知道，若$a=1$，則

$x_{n+1}=(x_n+c)\bmod m=(x_0+nc)\bmod m.$

是故不會有好的亂數數列產生。（如上例Ex2，所得是{0，1，2，3，4，5，6，7}）

由$x_{n+1}=(ax_n+c)\bmod m$可推得$x_{n+k}=\left(a^k x_n+(a^k-1)\frac{c}{b}\right)\bmod m$, ……②　　$k\geq 0$，$n\geq 0$，$b=a-1$，若$a\geq 2$

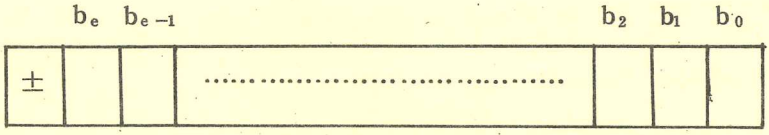
由②知：若 $n=0$ ，$\Rightarrow\ x_k=(a^k x_0+(a^k-1)\frac{c}{b})\bmod m, k\geq 0$ ，換句話說：$\{x_k \mid k\geq 0\}$ 都可由 x_0 ，a ，c ，m四個數來作決定，故L.C.M所產生的亂數，項項並非不相干，但是仔細的選取a ，c ，m ，所得的數列 $\{x_k\}$ 仍能有相當不錯的性質。（例如：週期長，能通過某些統計檢定）。

由①式知：所產生的 $\{x_n\}$ 週期最長爲m，爲了使對應的 $\{U_i\}$ 能均勻的分布在〔0 ，1 〕上；使 $\{U_i\}$ 接近U〔0 ，1 〕，則必須m足夠大，及週期夠長。由於m大，對應的 $\{U_i\}$ 才能更稠密於〔0 ，1 〕上。而從計算效率觀點來看：在計算機中除法是運算中較慢的一步驟，若能在計算①式時避開除法，則會增快執行的速度。下一段，我們就此一觀點來作如何適當選擇m的決定，至於如何使週期夠長，則稍後討論。

第二節 （上）

如何適當選擇m？

設 ω 是計算機中字組位字數（word size）。而e是字組（word）中除去正、負位元（i.e, $e=\omega-1$），例如：IBM360，370是32-bits/word故 $\omega=32$ ，而最左邊是正、負位元，$\therefore e=\omega-1=31$ 。通常m都是取 2^e；原因有2 ，首先是一般計算機中 2^e 通常是很大的數（Ex：$2^{31}>10^8$ 以上）；第二是可以避免除法：



原因是：此計算機能表示的非負整數是由〔$0\sim 2^e-1$〕，換句話說，若欲得 $B=A \bmod 2^e$ 。

case(i)：$A<2^e$ ，則B＝A ，i.e., B就是location A 中所現存的數。

case(ii)：$A\geq 2^e$ ，則當A值欲置入位址A是就已經發生整數溢位（integer overflow），其置入的結果，將使A值左邊位元

中超過 2^e 的部份損失掉。結果置入A位址中的值（損失過）就是mod 2^e 後的值。

我們舉一例子：設本計算機中一字組（word）是5 bits. i.e $\omega = 5$，$\Rightarrow b = 4$，取 $m = 2^b = 2^4$

case(i)：若位址A中置有12，i.e.,

location A

sign	1	1	0	0

則B = A mod m是12, i.e.,

location B

sign	1	1	0	0

case(ii)：若在某運算中，欲將結果27置於A位址中，然後再將27 mod m存到B，則其過程可寫：$27 = (\quad 11011)_2$

置入：A位址會發生integer orerflow，導至左邊超過4位元者之損失，換句話說A中所得者已經27 mod 16 = 11，故只要再B ← load A 即可。

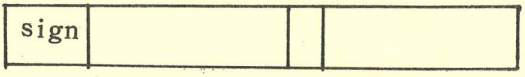
location A

sign	1	0	1	1

由上示case(i)、case(ii)可以很明顯的發現，根本不須要用除法，反正利用integer overflow的特性，即可得A中的值即寫A mod 2^e。又若欲置入A的值若爲負值，由於2的補數上的方便，置入後的值照樣可以視成mod後的值。所必須提醒的是：integer overflow及2's Complement的特性，並非所有的計算機都有。

由以上的討論，我們可以方便的演算法來完成 $x_{n+1} = (ax + c) \bmod 2^e$，設計算機ALU中的Register是 2ω bits i.e.，由一個 T_1 及 T_2（各是 ω bits）連接而成（似double precision）。

Regizter T_1 T_2



設 $A \leftarrow a$ ， $C \leftarrow c$ ， $X \leftarrow x_n$

則：Algorithn

①LOAD x to Register

②Multiply Register and A, store in Register.（暫存）

③LOAD T_2 to x（注意T_2的sign與T_1一樣）

④ADD x and c , store in Register.（暫存）

⑤LOAD T_2 to x

⑥output x

已經討論過m的取法後，讓我們再看看何時可以得到最大週期？

第二節 （中）

如何求得最大週期？

當然，如果只是要找一個有最大週期的數列，則 $x_{n+1} = (x_n + 1) \bmod m$ 即可提供。（$x_0 = 0$），但明顯的有的數列雖具有最大週期，但其性質不好，因此我們必須再尋找其他長週期的數列，我們必須擁有一套系統性的方法來尋找。以下的定理就是提供我們判斷的良好工具。

〔定理1〕 令 $\{x_i \mid i = 0, 1, \cdots\}$ 是由①所產生的數列

①m與c互質

本數列擁有全週期 λ $\Leftrightarrow$ ② $a \equiv 1 \bmod p$ ，p是m的質因數

（i.e., $\lambda = m$） ③若 $4 \mid m$ ，則 $4 \mid (a-1)$

（定理一）的證明是由Hull及Dobell在1962年完成；而最早的證明是由M. Greenberger證明了 $m = 2^e$ 的情況。

要證明本定理，我們先得證明二個Lemma ，以下證明主要參考〔Knuth〕。

《Lemma 1》 設p是質數，且e是一個正整數，$p^e > 2$ ，若 $x \equiv 1 \bmod p^e$ ，$x \not\equiv 1 \bmod (p^{e+1})$ ，則 $x^p \equiv 1 \bmod (p^{e+1})$ ，但 $x^p \not\equiv 1 \bmod (p^{e+2})$

【證明】 由條件知：$x \equiv 1 \bmod p^e$ i.e., $x - 1 = p^e q$ ，$q \in Z$（整數）

又 $x \not\equiv 1 \mod p^{e+1}$, i.e., , q 非p的倍數

$$x=1+p^e q \Rightarrow x^p=(1+p^e q)^p=1+\binom{p}{1}p^e q+$$

$$\binom{p}{2}(p^e q)^2+\cdots\cdots+\binom{p}{p-1}(p^e q)^{p-1}+(p^e q)^p$$

$$\Rightarrow x^p-1=(p^{e+1})\left[\frac{\binom{p}{1}}{p}q+\frac{\binom{p}{2}}{p}p^e q^2+\cdots\cdots+\right.$$

$$\left.\frac{\binom{p}{p-1}}{p}(p^e)^{p-2}q^{p-1}+\frac{p^{pe}q}{p}\right]$$

$\binom{p}{k}$ 是二項式係數，故能被 p 整除

$$=(p^{e+1})[q+(s)]$$

$$=p^{e+1}[q+p\overline{s}], \overline{s}=\frac{(s)}{p}$$

(s) = 是能被整除的整數。

$$=\binom{p}{2}p^{e-1}q^2+\cdots+\binom{p}{p-1}p^{(p-2)e-1}q^{p-1}+p^{pe-1}q$$

$\therefore x^p-1\equiv 0 \mod p^{e+1}$, or $x^p\equiv 1 \mod p^{e+1}$

但因q不能被p整除 $\therefore [q+p\overline{s}]$ 無法再提出一個 p 的因數

故 $x^p-1 \not\equiv 0 \mod (p^{e+2})$ i.e.,

$x^p \not\equiv 1 \pmod{p^{e+2}}$

《Lemma 2》 若將m分解成質因數的乘積 i.e., $m=p_1^{e_1}p_2^{e_2}\cdots\cdots p_t^{e_t}$, 若 λ 是由 (x_0, a, c, m) 所產生數列的週期，而 λ_j 是由：$(x_0 \mod p_j^{e_j}, a \mod p_j^{e_j}, c \mod p_j^{e_j}, p_j^{e_j})$ 所產生數列的週期 $1 \le j \le t$ ，則

$\lambda = \lambda_1, \lambda_2, \cdots\cdots, \lambda_t$ 的最小公倍數。

【證明】 本Lemma的證明是採用對大作歸納法。我們只要證明若 m_1 , m_2 互質，則由 (x_0, a, c, m_1m_2) 所生數列週期 λ_m 是 $(x_0 \mod m_1, a \mod m_1, c \mod m_1, m_1)$ 及 $(x_0 \mod$

m_2 , a mod m_2 , c mod m_2 , m_2) 所生週期 λ_{m1} 及 λ_{m2} 的最小公位數， i.e., $\lambda_m = \text{lcm}(\lambda_{m1}, \lambda_{m2})$ 。

令 $(x_0, a, c, m_1 m_2)$, $(x_0 \bmod m_1, a \bmod m_1, c \bmod m_1, m_1)$, $(x_0 \bmod m_2, a \bmod m_2, c \bmod m_2, m_2)$

所產生的數列各爲 $\{\hat{x}_n\}$, $\{\hat{y}_n\}$, $\{z_n\}$ 。

由 $x_{n+1} = (a x_n + c) \bmod m_1 m_2$

可知： y_{n+1} 可視由 x_{n+1} 再取mod m_1 , Z_{n+1} 可視由 x_{n+1} 再取mod m_2

即 $y_n = x_n \bmod m_1$, $z_n = x_n \bmod m_2$

我們可以有： $x_k = x_n \Rightarrow y_k = y_n \quad z_k = Z_n$ ……③

設 $\lambda' = \ell\text{cm}(\lambda_{m1}, \lambda_{m2})$ ，則我們的目標在於 $\lambda' = \lambda_{m1 m2}$

因爲 $x_n = x_n + \lambda_{m1 m2}$ (對每一適當的n) ，

則由③知： $y_n = y_{n+\lambda_{m1 m2}}$ 及 $z_n = z_{n+\lambda_{m1 m2}}$

是故 $\lambda_{m1 m2}$ 是 λ_{m1} 及 λ_{m2} 的倍數 $\Rightarrow \lambda_{m1 m2} \geq \lambda'$ ……④

又由 $y_n = y_{n+\lambda'}$ 與 $z_n = z_{n+\lambda'}$ (對每一適當的n) ，

由③知： $x_n = x_{n+\lambda'}$

是故 λ' 是 $\lambda_{m1 m2}$ 的倍數 i.e., $\lambda' \geq \lambda_{m1 m2}$ ……⑤

由④及⑤知道 $\lambda_{m1 m2} = \lambda' = \ell\text{cm}(\lambda_{m1}, \lambda_{m2})$

完成此二Lemma 後，我們可以開始來證明定理一。

由Lemma 2 知道：若 $\lambda = \ell\text{cm}(\lambda_1, \lambda_2, \cdots\cdots, \lambda_t)$

$$\leq \lambda_1 \lambda_2 \cdots\cdots \lambda_t \leq p_1^{e1} p_2^{e2} \cdots\cdots p_t^{et}$$

顯然 $\lambda = m = p_1^{e1} p_2^{e2} \cdots\cdots p_t^{et}$ (i.e., 最大週期時)

$\Leftrightarrow \quad \lambda_j = p_j^{ej}$, $\forall_j$

〔註：任何整數，都可分解成有限個質因數的乘積〕

是故，我們可以假定 $m = p^e$ (i.e., 只有一個質因數)，如此，已減低很多複雜性。

當 $a = 1$ 時，由於 $x_n = (x_0 + nc) \bmod p^e$ ，則本數列週期 $\lambda = p^e$

$\Leftrightarrow$每個$x \in Z$，$0 \leq x < p^e$，都會出現一次。

$\Leftrightarrow$取$x_0 = 0$時，本數列的週期是p^e

$\Leftrightarrow x_n = (nc) \bmod p^e$，唯有$n = p^e$時，$x_n$才會再度爲0

$\Leftrightarrow$c與p互質。 是故$a = 1$時，本定理成立

當$a > 1$時：週期$\lambda = p^e$ $\Leftrightarrow$由$x_0 = 0$所產生的數列週期爲p^e

則由②式知：$x_n = \frac{a^n - 1}{a - 1} c \bmod p^e$ ……⑥

由⑥知：若c與m不爲互質，則x_n不產生1，〔因爲$\exists q \in Z$〕

$x_n = (\frac{a^n - 1}{a - 1} c) - p^e q$，由於產生最大週期$x_n$必有等於1

i.e., $(\frac{a^n - 1}{a - 1}) c - mq = 1$，for 適當的n，

i.e., $(c, k) = 1$

故若$\lambda = p^e$＝最大週期時，c與m互質，取$x_0 = 0 = x_\lambda = (\frac{a^\lambda - 1}{a - 1})c \bmod p^e$，i.e., $\frac{a^\lambda - 1}{a - 1} \equiv 0 \bmod p^e$。是故定理一中，$a > 1$的部份，我們已證明了：若數列最大週期 $\Rightarrow$ c與m互質，因此定理1的證明只差：若c與m互質，則數列有最大週期 $\Leftrightarrow \begin{cases} a \equiv 1 \bmod p & p > 2 \\ a \equiv 1 \bmod 4 & p = 2 \end{cases}$

而由上的討論知道：若$\lambda = p^e$最大週期 $\Rightarrow$ c與m互質 $\Rightarrow \frac{a^\lambda - 1}{a - 1} \equiv 0 \bmod p^e$，故我們只要再證明下面的Lemma。

由於λ是滿足$\frac{a^\lambda - 1}{a - 1} \equiv 0 \bmod p^e$，的最小正整數，再與⑥及$(c, m) = 1$的關係可知：Lemma 3中的$\lambda$即爲最大週期。

《Lemma 3》 設$1 < a < m = p^e$，假設λ是滿足$\frac{a^\lambda - 1}{a - 1} \equiv 0 \pmod{p^e}$的最小正整數，

則： $\lambda = p^e \Leftrightarrow \begin{cases} a \equiv 1 \bmod p & p > 2 \\ a \equiv 1 \bmod 4 & p = 2 \end{cases}$

【證明】 “⇒”

(1)設 $\lambda = p^e$ ，利用反證：若 $a \not\equiv 1 \mod p$ ，則 $\frac{a^n-1}{a-1} \equiv 0 \mod p^e \Leftrightarrow a^n - 1 \equiv 0 \mod p^e$ ，又由條件可知：$a^{p^e} - 1 \equiv 0 \mod p^e \Rightarrow a^{p^e} - 1 = p^e k_1$

$\therefore \quad a^{p^e} \equiv 1 \mod p \cdots\cdots ⑦$

又 p 是質數，∴由 Fermat 定理知道：

$a^{p^e} \equiv a \ (\mod p) \cdots\cdots ⑧$

由⑦與⑧發現 $a \equiv 1 \mod p$ 矛盾

(2)設 $p = 2$ ，且 $a \equiv 3 \mod 4$ ，$\Rightarrow \exists k \in Z \ni a = 3 + 4k$

$\therefore a^2 = 16k^2 + 24k + 8 + 1 \Rightarrow a^2 \equiv 1 \mod 8$ ，依照 Lemma 1 ，則可有 $a^4 \equiv 1 \mod 16$ ，$a^8 \equiv 1 \mod 32$ ，… $a^{2^{e-1}} \equiv 1 \mod (2^{k+1})$

$\Rightarrow a^{2^{e-1}} - 1 = 2^{e+1} k_1$ ，令 $z_1 = \frac{a^{2^{e-1}} - 1}{a-1}$

則 $(a-1) z_1 = 2^{e+1} k_1$ ，

但 $a \equiv 3 \mod 4 \quad \therefore a - 1 = 2 *$ (odd)，odd 是奇數

$\therefore z_1 = \frac{2^{e+1} q}{2 \times \text{odd}} = \frac{2^e q}{\text{odd}}$ ，但 $z_1 \in Z$ （整數）

$\therefore z_1 = 2^e \times E$ ，$E = \frac{q}{\text{odd}} \in Z$

是故，可有 $\frac{a^2-1}{a-1} \equiv 0 \mod 2^e$

∴依條件，$\lambda = 2^{e-1}$ ，矛盾

(2)－(2)若 $a \equiv 2 \mod 4 \Rightarrow a = 2(1 + 2k)$

而 $\frac{a^\lambda - 1}{a-1} = a^{\lambda-1} + a^{\lambda-2} + \cdots\cdots + a + 1 =$ 必是奇數

$\therefore \frac{a^\lambda - 1}{a-1} \not\equiv 0 \mod 2^e$

是故由(1)，(2)，(2)－(2)知，

若 $\lambda = p^e \Rightarrow \begin{cases} a \equiv 1 \mod p & p > 2 \\ a \equiv 1 \mod 4 & p = 2 \end{cases}$ 成立

"⇐"由 $\begin{cases} a \equiv 1 \mod p & p > 2 \\ a \equiv 1 \mod 4 & p = 2 \end{cases}$

知：a 可寫成：$a = 1 + qp^f$，$p^f > 2$，且 q 非 p 的倍數

則 $a \equiv 1 \mod p^f$，而 $a \equiv 1 \mod p^{f+1}$，應用Lemma1，可得：$a^{p^g} \equiv 1 \mod (p^{f+g})$，但是

$a^{p^g} \not\equiv 1 \mod (p^{f+g+1})$，$\forall\ g \geq 0$

是故：$a^{p^g} - 1 = (p^{f+g}) \cdot (k_1)$，$a - 1 = qp^f$；

$a^{p^g} - 1$ 可以提出 $(a - 1)$ 的因式

$$\therefore \frac{a^{p^g} - 1}{a - 1} = \frac{p^{f+g} \cdot k_1}{a - 1} = \frac{p^{f+g} \cdot k_1}{qp^f} \in Z$$

$$\therefore \frac{a^{p^g} - 1}{a - 1} \equiv \mod (p^g) \cdots\cdots ⑨$$

但是 $\frac{a^{p^g} - 1}{a - 1} \not\equiv 0 \pmod{p^{g+1}}$ ……⑩

取 $g = e$，則 $\frac{a^{p^e} - 1}{a - 1} \equiv 0 \mod (p^e)$。而由②式知：由 $(0, a, 1, p^e)$ 所決定的數列可有：$x_n \equiv \frac{a^n - 1}{a - 1} \mod p^e$，所以，由 λ 的定義知道（Lemma 3 的條件中），$\{x_n\}$ 具有週期 λ，即 $x_n = 0 \Leftrightarrow n$ 是 λ 的倍數，又 $x_{p^e} = 0$，$\therefore p^e$ 是 λ 的倍數。

又 p 是質數，則 λ 必是 p^g 形式（for some g），i.e., $\lambda = p^g$

而由Lemma 3 的條件知：$\frac{a^\lambda - 1}{a - 1} \equiv 0 \mod p^e$，

故 $\frac{a^{p^g} - 1}{a - 1} \equiv 0 \mod p^e$

而由⑨及⑩知：$g = e$，是故 $\lambda = p^e$ 。

我們應用已經討論過的，若選取的$m=2^e$，則利用本定理，即刻可獲得一最大週期的數列：只要取$a-1\equiv 0 \mod 4$，c爲奇數，即可達到目的。

第二節 （下）

我們已經詳細的看過如何選取m，a及c來獲得最大週期。同時，我們對$c=0$的狀況也有興趣——因①$c=0$是Lehmer最早提出LCM時所探討，②$c=0$的時候，執行的速度會比①式快！通常，我們稱$c=0$的情況爲：multiplicative congruential generator 而$c\neq 0$時，稱爲Mixed generator 。

當$c=0$時，①式變成：$x_{n+1}=ax_n \pmod m$ ……⑪

由定理一知道：$\{x_n\}$無法達成最大週期m，例如：$x_n\neq 0$，$\forall_n$，否則$\{x_{n+1}, x_{n+2}\cdots\cdots\}$都是0。而且若$(x_n, m)=d$，$\neq 1$時，則$x_{n+j}$，$j\geq 0$都將是d的倍數，不夠“亂”。所以，當$c=0$時，我們要求$x_n$與m互質$\forall_n \geq 0$，又最大的$x_n$可能小於或等於$m-1$，所以$\{x_n\}$就是那些小於m並且與m互質的非負整數。依照Euler's function φ的定義知：$\{x_n\}$的週期λ至多寫$\varphi(m)$。顯然，當m是質數時，$\varphi(m)=m-1$，故雖然$c=0$無法滿足定理一，而得最大週期m，但它有可能達到週期爲$m-1$，是故在此$m-1$的週期下，$\{x_n\}$能均勻分佈在$1\leq x_n \leq m-1$上。（$\because x_n\neq 0$）。

現在，讓我們來看看，要如何選取a及m，才能使$\{x_n\}$的週期儘量長！依照Lemma 2，我們仍可假設$m=p^e$，$\Rightarrow x_n=a^n x_0 \mod p^e$ ……⑫

由⑫，若a是p的倍數，則$\{x_n\}$的週期將很短，故(a, p)要互質。若設λ是$\{x_n\}$的週期$\Rightarrow x_0=a^\lambda x_0 \mod p^e \Rightarrow (a^\lambda-1)x_0=p^e q$，$q\in Z$……⑬

對於⑬，我們作以下的討論：

case(i)：若$(x_0, p^e)=p^f \Rightarrow a^\lambda \equiv 1 \mod (p^{e-f})$；而由Euler定理可知：$a^{\varphi(p^{e-f})}\equiv 1 \mod (p^{e-f})$，

$\therefore \lambda$是$\varphi(p^{e-f})$的因數，又$\varphi(p^{e-f})=p^{e-f-1}(p-1)$

$\therefore \lambda \mid p^{e-f-1}(p-1)$

case(ii)：若$(a, p^e)=(a, m)=1$，$\Rightarrow$滿足$a^\lambda \equiv 1 \mod(p^e)$的最小整數$\lambda$，就是$\{x_n\}$的週期，$\lambda$稱爲“order of a modulo m”。

case(iii)：若$(x_0, p^e)=1$，則由⑬得$a^\lambda - 1 = p^e q_1$，$q_1 = \frac{p}{x_0}$

$\therefore a^\lambda \equiv 1 \mod p^e \Rightarrow \lambda \mid p^{e-1}(p-1)$

若任何一個數，有最大的orderof modulo m，則此種數叫作：primitive element modulo m。而一個最大可能order modulo m，記作$\lambda(m)$，則是⑭中所述。歸納所論：當$c=0$時，$\lambda(m)$就是$\{x_n\}$所能達到的最大週期。

〔Knuth，p-19〕中舉出：

$$\lambda(2)=1,\ \lambda(4)=2,\ \lambda(2^e)=2^{e-2} \quad \text{if } e\geq 3$$

$$\lambda(p^e)=p^{e-1}(p-1) \quad \text{if } p>2 \qquad \cdots\cdots⑭$$

$$\lambda(p_1^{e_1}\cdots p_t^{e_t})=\ell cm(\lambda(p_1^{e_1}),\cdots,\lambda(p_t^{e_t}))$$

由上面討論及⑭可以知道：若取$m=2^e$，則其最大週期爲2^{e-2}，換句話說，縱然在其所能達的最大週期2^{e-2}，亦只有$m=2^e$的$\frac{1}{4}$，隨之而來的問題是：這$\frac{1}{4}$的整數到底如何分布，也許都擠在某一邊也說不定；是故我們必須再尋求解決之道。若m＝質數，則情況比取$m=2^e$時好多了，綜合前面所述，m該選取比2^e小的最大質數。

當$c=0$時，爲了達到最可能的長週期，我們必須選a爲primitive element mod m及m是質時。那我們現在面臨的問題是：

⑴如何選擇一個具有primitive element mod m性質的a？

⑵當m是質數，則我們仍能避免直接的除法嗎？

對於問題①，Knuth收集了當$m=p^e$時的狀況：

〔定理二〕 若a是整數，且是一個primitive element modulo p^e（p是質數）。

$\Leftrightarrow$(i) $p^e=2$且a是奇數，

或$p^e=8$且$a\equiv 3, 5, 7 \mod 8$

或$p^e = 4$且$a \equiv 3 \mod 4$或$p = 2$，$e > 4$，

$a \equiv 3, 5 \mod 8$

或：(ii) 若p是奇數，$e = 1$，$a \not\equiv 0 \mod p$且$a^{\frac{p-1}{q}} \equiv 1 \mod p$，其中q是$p-1$的任意質因數。

或：(iii) 若p是奇數，$e > 1$，$a \not\equiv 0 \mod p$，

$a^{p-1} \not\equiv 1 \mod p^2$

故若選$m = 2^e$，則a只要選$a \equiv 3$ or $5 \mod 2^e$即可，（注意此時只能週期達2^e的$\frac{1}{4}$），若$m =$質數$p > 2$，則須滿足(ii)。

對於問題②，由於m是小於2^e的最大質數，令$m = 2^e - q$，for some q。

$$\Rightarrow x_{n+1} = a x_n \mod m = a x_n \mod (2^e - q)$$

$$\therefore x_{n+1} = a x_n - (2^e - q)k, \text{ for some } k。$$

$$\Rightarrow x_{n+1} = a x_n - 2^e k + qk$$

$$= a x_n \mod 2^e + qk$$

我們令 $\overline{x}_{n+1} = a x_n \mod 2^e$，又k爲$\text{INT}(\frac{a x_n}{2^b - q})$是故：

$$\text{若 } \overline{x}_{n+1} + qN < 2^b - q \Rightarrow x_{n+1} = \overline{x}_{n+1} + qN$$

$$\overline{x}_{n+1} + qN > 2^b - q \Rightarrow x_{n+1} = \overline{x}_n + qN - (2^b - q) \quad \cdots\cdots⑮$$

由上面論述可知：要求$\overline{x}_{n+1}$仍然可以使用integer overflow的方式，避免了除法，此種技巧是由Payne，Rabung與Bogyo所提出，並稱爲模擬除法（simulated division）。

應用simulated division編製一個與機型無關（machine independent）的程式是可以辦到的。例如若$e = 31$（如I.B.M.360/370），則最大質數恰好是$2^{31} - 1$，而primitive element a取$a = 7^5 = 16807$，則要執行$x_{n+1} = 7^5 \ x_n \mod (2^{31} - 1)$，利用simulated division的方式，得如下的方式：

$$x_{n+1} = 7^5 \ x_n \mod (2^{31} - 1) = (7^5 \ x_n \mod 2^{31}) + k \quad \cdots\cdots⑯$$

須注意的是：2^{31}，已經超過計算機的最大正整數容量，所以處理

$\overline{x}_{n+1} = 7^5 x_n \bmod 2^{31}$ 時，須要用一些技巧。令 $a = 7^5$

由於 $\overline{x}_{n+1}$ 是 ax_n 除以 2^{31} 後的餘數，故只要把 ax_n 表成

$$ax_n = 2^{31} * \text{Quotient} + \text{Remainer}$$

$$= 2^{16} \cdot 2^{15} \cdot \text{Quotient} + \text{Remainder}$$

令 INT(A) = 〔A〕，i.e.，INT(A) 是小於等於A的最大整數。

$$\Rightarrow x_n = 2^{16} * \text{Quotl} + \text{Remainl}\ ,\ \text{Quotl}\ \ \text{INT}\left(\frac{x_n}{2^{16}}\right)$$

$$\text{Remainl} = x_n - 2^{16} * \text{Quotl} = x_n - 2^{16} * \text{INT}\left(\frac{x_n}{2^{16}}\right)$$

$$\therefore ax_n = 2^{16} * (a * \text{Quotl}) + a * \text{Remainl}. \quad \cdots\cdots ⑰$$

也許 $a * \text{Remainl}$ 會超過 2^{16}，故若令 $\text{TEMP} = a * \text{Remainl}$

$$\cdots\cdots ⑱$$

$$\Rightarrow \quad \text{TEMP} = \text{INT}\left(\frac{\text{TEMP}}{2^{16}}\right)\ \ 2^{16} + \text{Remain 2} \quad \cdots\cdots ⑲$$

由⑰，⑱，⑲知：

$$ax_n = 2^{16} * 〔a * \text{Quotl} + \text{INT}\left(\frac{\text{TEMP}}{2^{16}}\right)〕 + \text{Remain 2},$$

$$\text{where Remain 2} = \text{TEMP} - \text{INT}\left(\frac{\text{TEMP}}{2^{16}}\right) * 2^{16} \quad \cdots\cdots ⑳$$

令 $\text{EXTRA} = a * \text{Quotl} + \text{INT}\left(\frac{\text{TEMP}}{2^{16}}\right)$，

則 $ax_n = 2^{16} \cdot \text{EXTRA} + \text{Remain 2}$ $\cdots\cdots ㉑$

但是 $\text{EXTRA} = \text{INT}\left(\frac{\text{EXTRA}}{2^{15}}\right) * 2^{15} + \text{Remain 3},$

$$\text{where Remain 3} = \text{EXTRA} - \text{INT}\left(\frac{\text{EXTRA}}{2^{15}}\right)\ \ 2^{15}$$

則 $ax_n = 2^{16} * \text{EXTRA} + \text{Remain 2}$

$$= 2^{16}\ \ 2^{15}\left(\text{INT}\left(\frac{\text{EXTRA}}{2^{15}}\right)\right) + 2^{16} \cdot \text{Remain 3} +$$

Remain 2.

$$= 2^{31} * (\,INT(\frac{EXTRA}{2^{15}})) + 2^{16} \cdot Remain\ 3 + Remain\ 2$$

所以 $Quotient = INT(\frac{EXTRA}{2^{15}}) = k$

$$Remainder = 2^{16} \cdot Remain\ 3 + Remain\ 2$$

$$\therefore \overline{x}_{n+1} = Remainder = 2^{16} \cdot Remain\ 3 + Remain\ 2$$

而由⑮知：我們可先令：

$$x = \overline{x}_n + k - (2^{31} - 1)$$

若 $x < 0$，則用 $x \leftarrow x + (2^{31} - 1)$ 來修正即可！

第二節 （末）

我們以二個例子來做爲 L.C.M. 的結束。

〔例一〕 當 $c = 0$ 時，我們曉得只有選一 primitive element modulo $m = 2^e$ 即可獲致一個相當於 $\frac{1}{4}m$ 的週期。然而我們也曾提出懷疑到底這 $\frac{1}{4}m$ 的亂數，是否能均勻分佈於（0，m）間呢？我們由以下的例子就可知道：不同的 a，將對應到不同的分佈狀況！由定理二我們知道，當 $m = 2^e$ 時，a 可選取 $a \equiv 3$ or $5 \mod 8$，是故當 $e = 12$ 時，ie $m = 4096$，因理論知道，將產生 1024 個亂數。若令 $\bar{x} = x_n$，$y = x_{n+1}$，$n \geq 0$ 則理論上若 $\{x_n\}$ 能均勻分佈於（0，m），則（x，y）能均勻分佈於（0，m）×（0，m）的平面圖上。以下，我們令 $a = 3$，$a = 27$，$a = 61$，$a = 125$，並把（x，y）的圖畫出來：如下圖一～四。顯然在圖一上（$\bar{x}$，y）幾乎分佈於三條線上，現出相當壞的效果。

圖二，圖三，圖四分佈的較圖一地勻。值得注意的是圖 ϕ = 上的點分佈在 125 條線上，（圖 3161 條線上（圖二）27 條線上。（注意坐標軸的方向）。由 $x_n = ax_{n-1} \mod m$ 中，可看出來，$x_n = ax_{n-1}$ 是直線，是故 a 若選大，則經 mod 後連續的對，不

$x_{n+1} = 3\ x_n \mod 4096$

with $x_0 = 1$,

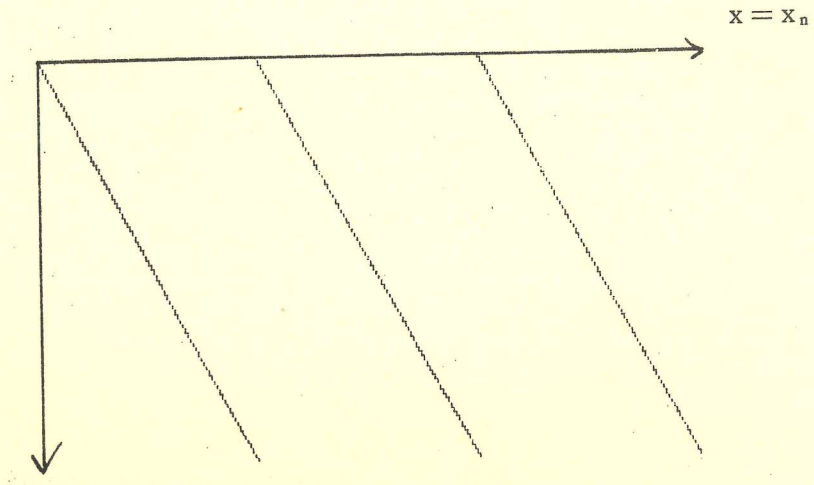


圖 一

$x_{n+1} = 125\ x_n \mod 2^{12}$

with $x_0 = 1$

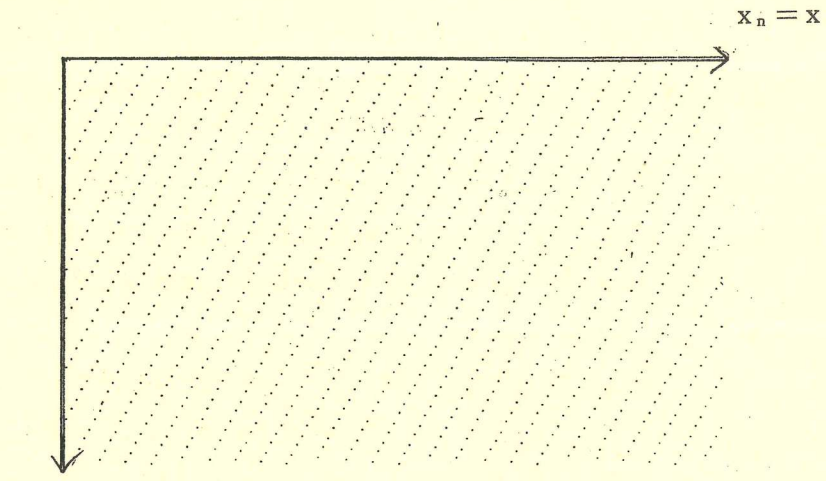


圖 二

$x_{n+1} = 61\ x_n \bmod 2^{12}$

with $x_0 = 1$

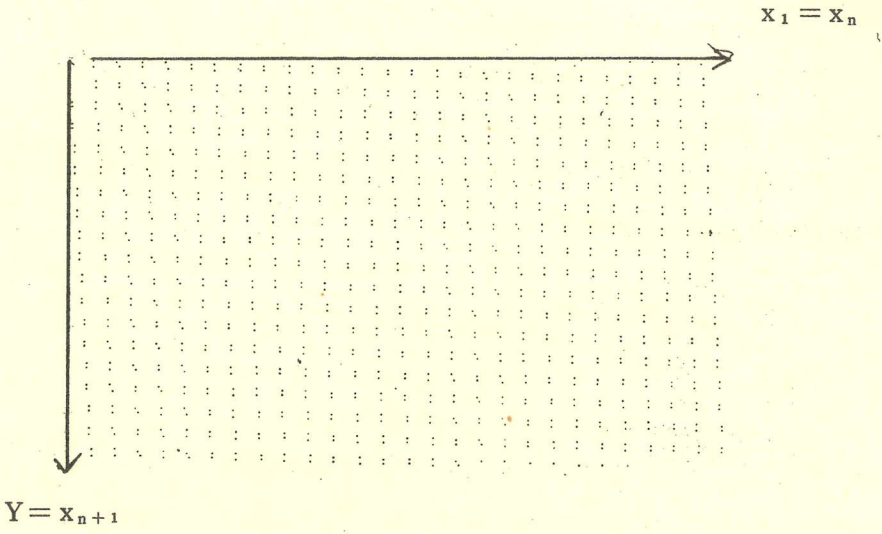


圖 三

$x_{n+1} = 27\ x_n \bmod 2^{12}$

with $x_0 = 1$

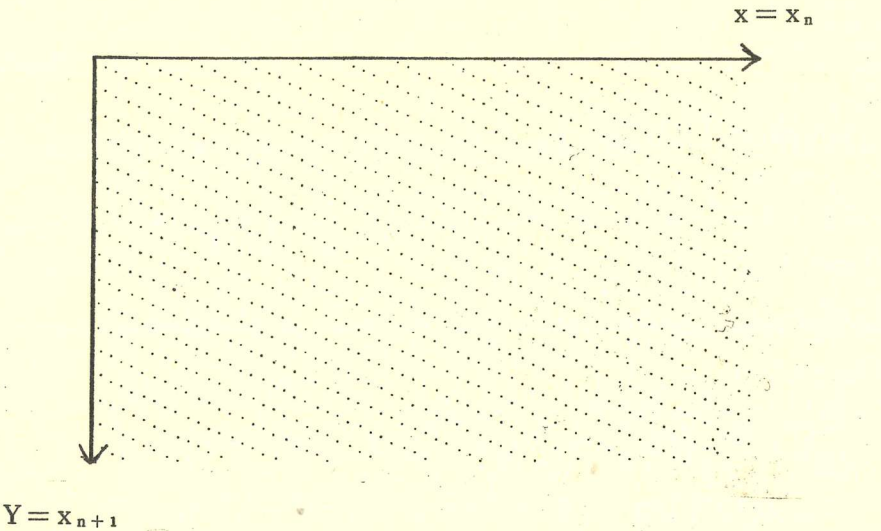


圖 四

致於分佈在同一條直線。例若a＝3，⇒（1，3），（3，9）是在同一直線上，但若是a＝2501，則（1，2501），（2501，9），不在同一直線上。

我們也可以討論3-dimensional的情況：i.e.，選取（x_n，x_{n+1}，x_{n+2}），看看它們是否均勻分布在（0，m）×（0，m）×（0，m）上。我們必須記住的是能在2-dimensional上表現好的a，不一定在3-dimensional上有好的表現：

例如：$x_n = 65539\, x_{n-1} \bmod 2^{31}$（曾爲IBM使用），在2-dimensional上表現好，但在3-dimesional時；考慮x_n，x_{n+1}，x_{n+2}的關係：

$$\begin{aligned} x_{n+2} &= 65539\, x_{n+1} \bmod 2^{31} \\ &= 65539\,(65539\, x_n) \bmod 2^{31} \\ &= 65539^2 x_n \bmod 2^{31} \\ &= (2^{16}+3)^2 x_n \bmod 2^{31} \\ &= [6\, x_{n+1} - 9\, x_n] \bmod 2^{31} \end{aligned}$$

是故（x_{n+2}，x_{n+1}，x_n），易於停留在平面$x_{n+2} = 6\, x_{n+1} - 9\, x_n$（∵6與9比起$2^{31}$太小了）。

〔例二〕 早期的研究者對於LCM的運算數率非常重視，想出若$a = 2^e + 1$，則$a\, x_n$的乘法可以“移位與加法”（shift and add）的方式來代替。注意當$m = 2^e$時，a的選法符合定理一。我們來看看如何不用乘法的運算來達成乘法的效果。

$x_n = (a x_{n-1} + c) \bmod m$，茲討論$a x_{n-1}$：

$a x_{n-1} = (2^{\ell} + 1)\, x_{n-1} = 2^{\ell} \cdot x_{n-1} + x_{n-1}$；而$2^{\ell} \cdot x_{n-1}$的效果是將$x_{n-1}$往左邊移$\ell$－bit，∴$a x_{n-1}$的結果可由：$x_{n-1}$往左移與$x_{n-1}$相加而得，避免了乘法。但是，最近的研究指出shift and add方式不見得好，因爲此種$a = 2^{\ell} + 1$通常有不好的統計性質。Knuth提出了“potency”的方式可以判斷之！

若 $x_{n+1}=(ax_n+c) \bmod m$ 能產生週期 m 的數列，則定義 $\{x_n\}$ 的 potency 是 $b^s \equiv 0 \bmod m$，$b=a-1$，$s \geq 1$ & 若 $s_1 < s \Rightarrow b^{s_1} \neq 0 \bmod m$. 應用定理的必要條件②知，此種 s 必定存在。爲了方便起見，令 $x_0=0$，由②式可知：

$$x_n=(a^n-1)c/b \bmod m.$$

由 $a^n-1=(b+1)^n-1 \Rightarrow x_n=c\left(n+\binom{n}{2}b+\cdots+\binom{n}{s}b^{s-1}\right) \bmod m.$ ……㉒

若 $a=1$，則 potency $=1 \Rightarrow x_n \equiv cn \pmod m$ 一點也不亂。

若 potency $=2$，則 $x_n=cn+cb\binom{n}{2} \Rightarrow x_{n+1}-x_n \equiv c+cbn.$ 也不亂。

若考慮 3-dimensional (x_n, x_{n+1}, x_{n+2})，則其易布於：

$x-2y+z=d+km$ 上，$k=0, 1, -1, -2,$

$d=cb \bmod m.$

〔Knuth〕提及，好的亂數數列其 potency 大都要 ≥ 5。

現在我們回頭來看一個：$a=2^k+1$ 型的例子（設 $m=2^{35}$）。

$\because a-1=2^k \therefore b=2^k \Rightarrow$ 若 $k \geq 18$ 時，$b^2 \equiv 0 \bmod m$.

是故 potency $=2$

$\Rightarrow$ 若 $k=12, 13, \cdots 17$，$b^3 \equiv 0 \bmod m$

是故 potency $=3$

唯有 $k \leq 8$ 時，potency 才能 ≥ 5，但是 $k \leq 8$，則 $a \leq 257$。

一般而言，a 不應太小，故在 $m=2^{35}$ 時，2^k+1 型的 a 應不予考慮。

雖然 potency 的觀念簡單，但是：Randomness $\Rightarrow$ high potency，high potency 不一定是 Randomness。

第三節　二次同餘法（Quadratic Congruential Method）（QCM）

由於 LCM 的結構很簡單，於是有人希望能夠擴充 LCM，來得到更具有"亂"（Randomness）的性質。本節及下面幾節將朝這個目標來說明。

首先我們看 $x_{n+1}=(dx_n^2+ax_n+c) \bmod m.$ ……㉓

本型式就稱爲：二次同餘法（Q.C.M），顯然的若是 $d \equiv 0 \mod m$，則㉓就回到①式！㉓式的週期最大可達到m，與①相同，但是 x_{n+1} 却只由 x_n 來決定，與 $\{ x_{n-1}, \cdots\cdots, x_0 \}$ 無關。

R.R. Coveyou 提出了一個類似於Midsquare Method. 的Q.C.M，但是，經證明具有相當長的週期及相當“亂”的效果。其型式如下：

令 $m = 2^e$，且 $x_0 \mod 4 = 2 \Rightarrow x_{n+1} = x_n^2 + x_n \mod 2^e$，

$n \geq 0$ ……㉔

我們可以將㉔式改變一下：$x_{n+1} = x_n^2 \mod 2^e + x_n \mod 2^e$ …㉕

然後再修正如下：若 $x_{n+1} > 2^e$ 則 $x_{n+1} \leftarrow x_{n+1} - 2^e$

雖然㉔式比①式複雜，但我們仍能利用①式計算上的技巧來處理㉔式，我們以下列的演算法來說明：㉕式

設 Register R1 是 Double precision，i.e，R1 是由T1及T2連接而得。

設 x_n 的資料放在x處，而R2是另外一個 single precision的Registor。

則：①Load：R1 ← x

②Multiply：R1 and x，then store in R1〔i.e R1中是放 x_n^2〕

③Load：R2 ← T2（注意T2是R1的Low part，是故T2中的資料，即爲 $x_n^2 \mod 2^e$）

④SuB：$x - (2^e - 1)$，then store in x

⑤ADD：R2 + x，and then store in x

⑥SuB：x ← x − 1

⑦Branch：IF Rx > 0 then exit.

⑧ADD：x ← x + $(2^e - 1)$，then store in x.

⑨ADD：x ← x + 1

結果放在x處。

前面曾提及㉔像Midsquare Method，我們在此作一說明，並做爲本節

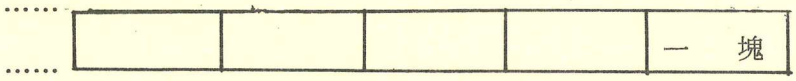
的結束。

假設一個word有5e bit以上

若$Y_n = 2^e x_n \Rightarrow Y_{n+1}$可由$Y_n^2 + 2^e Y_n$的中間2e－bit獲得。

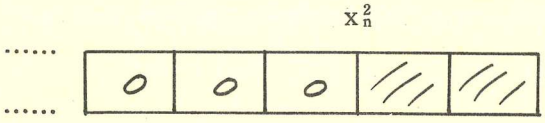
$\forall\ n \geq 0$，$x_n \leq 2^e$，$\Rightarrow x_n^2 \leq 2^e \cdot 2^e$，我們把字元word分塊：

每e－bits構成一塊：

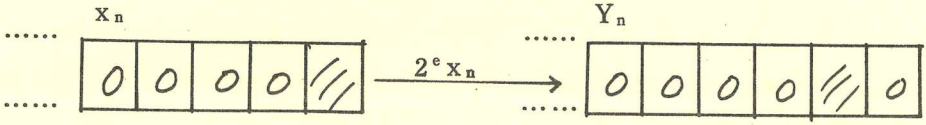


e—bits

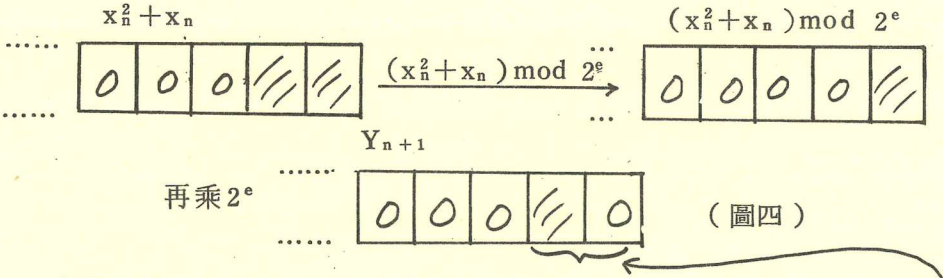
則x_n^2至多佔有二塊：



又$Y_n = 2^e x_n$，而x_n頂點佔一塊，再乘上2^e的效果即爲：將x_n往左移一塊！

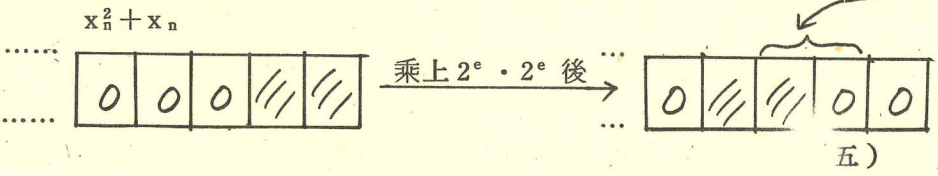


由定義知：$Y_{n+1} = 2^e x_{n+1} = 2^e ((x_n^2 + x_n) \bmod 2^e)$



（圖四）

但是：$Y_n^2 + 2^e Y_n = (2^e x_n)^2 + 2^e (2^e x_n) = 2^e \cdot (2^e (x_n^2 + x_n))$，故：



五）

比較圖四及圖五，可知：Y_{n+1} 是 $Y_n^2 + 2^e Y_n$ 的中間二塊。

第四節

由於微電腦普遍，我們是否可以應用它來產生亂數呢？若用LCM 則其週期最長才 m 。所以是否能發展出一種容易的方法，使得週期超過m，這一點很重要，若能達成此一目標，則我們也可在微電腦上產生效果不錯的亂數了。首先，我們看一個簡單的例子：若令 $x_n = (x_{n-1} + x_{n-2}) \bmod m$，則此種數列週期最長可達到$m^2$〔Ex：$x_0 = 1$，$x_1 = 2$，$m = 15$，⇒得{ 1，2，3，5，8，13，6，4，10，14，9，8，2，10，12，7，4，11，0，11，11，7，3，10，13，8，6，14，5，4，9，13，7，5，12，2，14，1，0，1，1，2 }，週期爲 40〕，顯然已達我們的願望——週期的增長。但是長週期只是代表著好的性質之一，但不幸地，由統計檢定觀點來看，本例 $x_n = (x_{n-1} + x_{n-2}) \bmod m$ 並非是好的，但這給了我們一項啟示：要得到更長（超過 m）並非不可能。

〔註：記得Fibonacii 數也是 $F_n = F_{n-1} + F_{n-2}$，故我們稱此例叫 Fibonacci sequence，在1950代早期曾被探討〕。

以下，我們稱 $x_n = (x_{n-1} + x_{n-k}) \bmod m$ 叫 Additive number generator，k 相當大。

1958 年時，G.J. Mitchell 與D.P. Moore 提出：

$$x_n = (x_{n-24} + x_{n-55}) \bmod m,\ n \geq 55 \quad \cdots\cdots ㉖$$

m是偶數，而 x_0，……，x_{54} 是任意不全爲偶數的整數。

本 additive generator 經證明其週期超過 $2^{55} - 1$，$m = 2^e$；是我們看過的最長週期！

㉖看起來不容易編碼，但Knuth 提出一個很有效率的演算法：（如后）

《Algorithm 1》 設Y〔1〕，Y〔2〕，…，Y〔55〕最初設初值 x_{54}，x_{53}，…，x_1，x_0，j ← 24，k ← 55，演算法中用簡單（具效率）的方式，重複使用Y〔1〕到Y〔55〕。

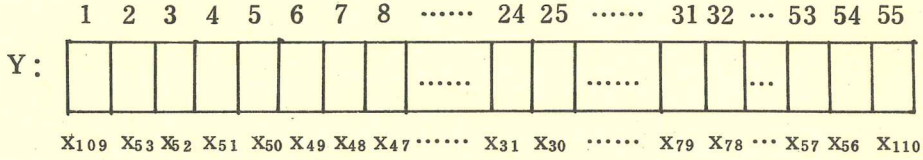
step 1. Y〔k〕← (Y〔k〕+Y〔j〕) mod 2^e

step 2 ，輸出Y〔k〕

step 3 ：j ← j －1 ，k ← k－1 ，若j＝0 則

j ← 55 ，若k＝0 則k ← 55　Go To step 1

我們稍作解釋，以便更清楚地了解Algorithm 1 ：



顯然x_0 用來產生x_{55} 後，就沒用處了，但x_{55} 日後會被用到，是故將x_{55} 放入Y（55）中，〔$x_{55} \leftarrow x_{31} + x_0 = Y〔24〕+Y〔55〕$，故初值j ← 24 ，k ← 55 ，而$x_{56} \leftarrow x_{32} + x_1 = Y〔23〕+Y〔54〕$，故有j ← j－1 ，k ← k－1的準備，而$x_{78} \leftarrow x_{54} + x_{23} = Y〔1〕+Y〔32〕$，此時j＝1 ，k＝32 。

下一步$x_{79} \leftarrow x_{55} + x_{24} = Y〔0〕+Y〔31〕$，i.e j＝0 ，但$x_{55}$ 放在Y〔55〕處，是故有若j＝0 時，j ← 55 。

⋮

$x_{109} \leftarrow x_{85} + x_{54} = Y〔25〕+Y〔1〕$，i.e., j＝25 ，但k＝1

而$x_{109} \rightarrow Y〔1〕$，又j ← j－1 ，k ← k－1 ∴j＝24 ，k＝0

$x_{110} \leftarrow x_{86} + x_{55} = Y〔j〕+Y〔k〕= Y〔24〕+Y〔0〕$，$x_{55}$

應在Y〔55〕處，故有：若k＝0 ，則k ← 55 ＝Y〔24〕+Y〔

Y〔55〕，而$x_{110} \rightarrow Y〔55〕$ ……

值得注意的是：本演算法的速度比以前所提的任一generator 都快——不但沒有除法，而且連乘法也省了。除了速度以外，它又是前所未見，具有超大週期，但是㉖未被廣泛推薦的原因在於：對於所產生的sequence 是否Randomness的研究過少，已致於我們了解它的只是：週期長、速度可快。

第五節　Feedback shift Register Methods（FSR）

Linear Congruential Method在較高維度上可能不比FSR分佈來的

均勻，此乃我們探究FSR原因！

1965年，Tausworthe介紹了下式（5-1）：

若$\{x_k\}$是由0 or 1所組成的序列，且x_k定義成：

$$x_k = (c_p x_{k-p} + c_{p-1} x_{k-p+1} + \cdots + c_1 x_{k-1}) \bmod 2 \quad \cdots\cdots ㉗$$

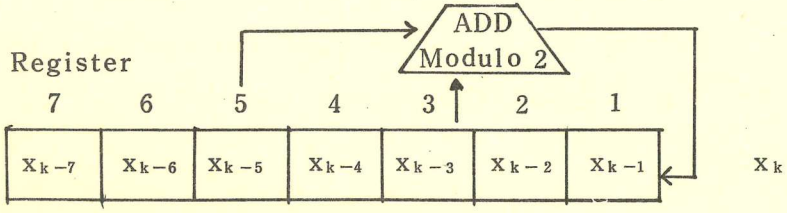
$\{c_p, c_{p-1}, \cdots, c_1\}$是固定一組0 or 1的常數。

則：$(z_{k-1}, z_{k-2}, \cdots, x_{k-p})$可視成一個連續回饋移位記錄器（shift regoster with feedback）。

（例5-1）

例如：令$p=7$，$c_5=c_3=1$，$c_1=c_2=c_4=c_6=c_7=0$

則：



初值	0	0	1	0	1	0	0	0
	0	1	0	1	0	0	0	0
	1	0	1	0	0	0	0	1
	0	1	0	0	0	0	1	0
	1	0	0	0	0	1	0	0

etc.

說明：$x_k = (x_{k-5} + x_{k-3}) \bmod 2$.

現將Register固定，而令register中編號3及5爲發生（ADD & mod 2）的地方，當運算開始後，可得到上示流動路線，而最右行（x_k下方）所得者$\{x_k \mid k > p\}$，爲所要的結果。

由於㉗可視成是一個具有p個分量i.e., $x_{k-1}, \cdots, x_{k-p}$的記錄器Regoster，而每一個x_i，$k-p \le i \le k-i$都有2種可能值（0或1）。

故$\{x_n\}$的週期最大可得2^p-1，（必須扣去（0，0，…，0），否則一旦出現（0，0，…，0），那以下的子數列都是（0，0，…，0））；

Tausworthe 在1965證明了要得最大週期2^p-1 $\Leftrightarrow f(x)=1+c_1x+\cdots\cdots+c_px^p$ ，$cp=1$ ，是佈於Galois field上，不可再分解的多項式，而此處Galois field是由{ 0，1 }組成，記成GF（2）；GF（2）上的運算是：一般的加法及乘法mod 2 。

（例5-2）

例子：（Leuis及payne所給）

令$p=5$ ，$c_1=c_2=c_4=0$ ，$c_3=c_5=1$ $\Rightarrow$

$$x_n=(x_{n-3}+x_{n-5})\bmod 2 \quad n\geq 5$$

但x_0 ，x_1 ，x_2 ，x_3 ，x_4 不全爲0 ，

令$x_0=x_1=x_2=x_3=x_4=1$ ，$\Rightarrow$

則前42個{ x_k }爲：111110001101110101000010010110011111000110

顯然本週期是$2^p-1=2^5-1=31$

由㉗所產生的{ x_k }是0與1所組成，那我們如何將此二元數列化成一般性的亂數數列呢？我們可取ℓ個bits爲一個亂數：如，在例（5-2）中所得是1111100011011101010000100101100 ，我們可取L＝4寫一組，得（15，8，13，13，4，2，5，9，15，1 ……。一般而言ℓ都相當接近計算機的word size，值得注意的是：本方法可以得到長週期，而且，與計算機的word size 無關！

第六節

有些人可能覺得Linear congruntial method or additive geneator 的方式太過容易，而無法給出足夠“亂”的隨機數，於是嘗試去合併二個generator來獲得第三個generator ，並且對於第三個generator所得之數列寄予厚望——期望它比前二個generator 所產生的數更亂。

假設我們已有了二個序列x_0 ，x_1 ，…，及Y_0 ，Y_1 ，…，介於0及m－1之間，最好（x_n）與（Y_n）無關。令$Z_n=(x_n+Y_n)\bmod m$；則若（x_n）的週期λ_1及（Y_n）的週期λ_2是互質時，$\Rightarrow$（Z_n）的週期是$\lambda_1\lambda_2$，相當長的週期。

〔證明〕 設 λ 是 (Z_n) 的週期，

則 $Z_n = Z_n + \lambda = (X_n + \lambda + Y_n + \lambda) \bmod m$.

又 (x_n)，(Y_n) 的週期分別是 λ_1，λ_2，i.e.,

$x_n + \lambda_1 = x_n$，$Y_n + \lambda_2 = Y_n$

$\therefore \lambda_1 \mid \lambda$，$\lambda_1 \mid \lambda \Rightarrow \lambda$ 是 $\lambda_1 \lambda_2$ 的倍數……①

又 $Z_n + \lambda_1 \lambda_2 = (x_n + \lambda_1 \lambda_2 + Y_n + \lambda_1 \lambda_2) \bmod m$

$= (x_n + Y_n) \bmod m = Z_n$

是故 $\lambda_1 \lambda_2$ 是 λ 的倍數 $\Rightarrow$ ……②

由①，②知： $\lambda = \lambda_1 \lambda_2$

另一種不同的技術由Maclaren 及Marsaglia (1965) 首先提出：

設給兩個不同的generator X，Y（產生的隨機序列是 (x_n) 及 (Y_n)），另設一個輔助用的表T：T〔0〕，T〔1〕…T〔k－1〕，（通常k是100左右，而最早提出時是k＝128），T的初值是令成 $\{x_0, \cdots, x_{k-1}\}$。則：

（步一）：由generator 分別產生X 及Y

（步二）：令 j＝INT〔k＊Y／m〕，則 $0 \le j \le k$，m是產生Y時，所用的modulo 。

（步三）：輸出T〔j〕，而再將T〔j〕← x

我們可以看出來：Y的作用是幫x 產生一“亂”的次序。Gebhardt 在1967時，曾證明此種技術所得到的新結果比原來 (x_n) 要更“亂”。然而，編成程式是却要注意（步二）：由於m都取得相當大，故k ＊ Y 有可能overflow由方法中可知，取二個generator 來產生結果；但（1976年）Carter Bays 及S.D. Durham將它改良了。

〔改良〕 只用原來的generator 就可以達到不錯的效果：

給一個generator x（所產生是 (x_n)），設另一輔助用的表T：T〔0〕，T〔1〕，…，T〔k－1〕，且T中初值設爲 $(x_0, \cdots, x_{k-1})$，又令 $Y = x_k$ 則：

〔步驟1〕 令 j ← INT〔k・Y／m〕，m是 (x_n) 中所用的

modulous . $\Rightarrow 0 \leq j < k$

〔步驟 2〕 Y←〔j〕，輸出 Y，再產生下（x_n）的下一項，存於〔j〕。

顯然，改良後的方式比原來的運算速度要快；而由過程可推知：它可能使新的（x_n）（經 Y，j 作用後）更“亂”。

第三章 統計檢定（Chi-square Test for Good of Fit）

在第二章中，我們看過了幾種主要產生亂數的方法，而所述的各種方式中，其 f(x) 都是固定的（如：LCM：$f(x)=(ax+c) \bmod m$），所以我們只能希望所產生的數列愈“亂”愈好（例如：$\{x_i\}$，對應的 $\{u_i\}$ 儘可能的接近 U(0,1)）。

因此，我們不僅要週期長，而且要有一套系統性的方式來做爲一個數列是否“亂”的評量工具。統計理論正好應運而來，有許多種檢定可以用來測量數列“亂”的程度。本章中，我們只看看如何用 Chi-Square Test 來測量已產生的數列。而我們須提醒的是：假設一數列已通過 T_1，T_2，T_n 種檢定，我們並不能保證它會通過另一個 T_{n+1}，我們所得到的是：能通過愈多檢定，則我們可以對它“亂”的程度更具信心。想參考其他的檢定方法者，可參閱〔Knuth〕。

在開始討論如何用 Chi-Squace Test 來檢定前，我們先對“good of Fit test”做一個概略性的介紹：假定我們對於某個隨機變數（Randam variable）的機率分配（probability distribution）未知，而我們想透過一些關於此項隨機變數之觀察資料（樣本）來判斷此項隨機變數的機率分配是否爲某種已假定的機率分配。例如：“我們想要判斷國中學生數學成績是否爲常態分配”，則我們抽樣幾所國中來得到樣本資料，然後再與已假定的機率分配：（本例中爲常態分配），比較看看是否假定的分配能“適合”（fit）於本樣本資料。

（註：最早且最有名的 goodness-of-fit test 是 Pearson 在 1900 年所提的 Chi-Square 檢定。）

現在，我們來看看如何使用Chi-Square 檢定。設由某個方法產生的亂數有n 個，我們把{ x_1 , x_2 , … , x_n } 化成對應的{ u_1 , u_2 , … , u_n } 。然後，將〔0 , 1〕等距分成k 個區間，我們希望{ u_1 , u_2 , … , u_n } 能均勻分布於〔 0 , 1 〕上。(一般而言，k至少須100 ，而$\frac{n}{k}$至少應該是 5) 。設F(x)是存在但未知的機率分配，則我們可依下列步驟來做檢定：

假設(Hyperthses) ：

Null : Ho : F(x) =U (0 , 1) for all x

alternative : H_1 : F(x) ≠U (0 , 1)for at least x.

i.e., Ho ：所觀查隨機變數之機率分配是連續均勻分配(0 , 1)

Ho ：所觀查隨機變數之機率分配非U (0 , 1)

檢定統計量(Test statistic) ：令j = 1 , 2 , … , k , 我們已把樣本分成k 類，令E_j 表示第j 類之理論期望值，i.e., 在所假設的U (0 , 1) 條件下：

$\therefore E_j = p_j N = \frac{n}{k}$；$p_j$ 是落在第j 類的機率，而我們是將〔0 , 1〕作了k 等分。

令0_j 表示樣本資料中落在第j 類的個數。〔註：我們所得的資料是{ x_1 , x_2 , … , x_n } ，實際執行時，可以不用先求出對應的{ U_1 , U_2 , … , U_n } ，而直接由{ x_1 , x_2 , … , x_n } 來做分類，事實上由{ x_1 , x_2 , … , x_n } 與{ U_1 , U_2 , … , U_n } 可以視成同義(equivalence)。

則我們的檢定統計量：

$$T=\sum_{j=1}^{k}\frac{(0_j-E_j)^2}{E_j}=\sum_{j=1}^{k}\frac{0_j^2}{E_j}-n=\frac{k}{n}\sum_{j=1}^{k}0_j^2-n \quad \cdots\cdots ㉘$$

只要n (樣本數目) 夠大，則㉘中T的近似機率分配是：自由度爲(k − 1) 的Chi-Square分配。又設$\chi^2_{k-1,1-\alpha}$ 是Chi-Square中俱有自由度k − 1 ，而對於機率爲1 − α 的百分點(percentage points) ，(例如：$\chi^2_{100,0.005}=67.3276$) 參考〔1〕，p − 984 頁。

則：當$T > \chi^2_{k-1,1-\alpha}$ 時，拒絕Ho ，否則接受Ho 。

〔註：若自由度（k－1）相當大，則我們可由（k－1）$\{1-\frac{2}{9(k-1)}+Z_{1-\alpha}*[\frac{2}{9(k-1)}]^{\frac{1}{2}}\}^3$來逼近$\chi^2_{k-1,1-\alpha}$，（見Gentle & Kennedy. p－118），其中$Z_{1-\alpha}$是常態分配N（0，1）對應到機率值$1-\alpha$的百分點percentage point〕。

第四章　亂數在數值分析方面的應用

若我們想要計算$I=\int_a^b f(x)dx$，f（x）是一實數值函數，若I存在，則我們可利用所產生的亂數來計算I值。

令Y＝（b－a）f（x），x是均勻分配於（a，b）上的連續型隨機變數，當然Y也是隨機變數。則

$$E(Y)=E((b-a)f(x))=(b-a)E(f(x))$$

$$=(b-a)\int_a^b f(x)\cdot\frac{1}{b-a}dx=\int_a^b f(x)dx=I$$

是故，要計算I，就相當於計算Y的期望值。

我們可以用樣本平均數（sample mean）來估計E（Y）：

$$令\overline{Y}(n)=\frac{\sum_{i=1}^{n}Y_i}{n}=(b-a)\frac{\sum_{i=1}^{n}f(x_i)}{n},\cdots\cdots㉙$$

由㉙知：只要由U（a，b）上產生n個數（亂數）x_i，則我們可以估計$\overline{Y}(n)$，又當n充分大時，$\overline{Y}(n)$可以相當接近於E（Y）＝I。

例子：（本例由〔Law & Kelton〕；p－50頁錄下）

若想計算$I=\int_0^{\pi}\sin x\,dx$，則由㉙式着手可有：

n	10	20	40	80	160
$\overline{Y}(n)$	2.213	1.951	1.948	1.989	1.993

理論值是2

第五章　結論

由以上的討論，我們可以感覺到；想要造一個好的亂數數列，並不容易也相當複雜，其所含蓋的預備知識相當的廣泛。而我們所要提醒的是：使用計算機中心所提供的亂數時，要特別小心，我們至少要使用某個檢定來測試它，以免所得的結果不可靠。

本文的主要目的是想提供高中教材，一些關於亂數（或隨機號碼）的資料。如更想深入研究，則可參考〔Knuth〕。

參考資料：

1. Conover, W.J.：PRACTICAL NONPARAMETRIC STATISTICS, 1980 2ed.（華泰）。chap.2 & 4 。
2. Kennedy, Jr, W.J. & GENTLE, J.E.：STATISTICAL Computing 1980.（華泰）. chap.6 。
3. Knuth, D.E.：The art of computer programming. vol.2, 2ed. 1980.（台北圖書公司）。
4. Law, A.E. & Kelton, W.D.：Simulation Modeling and Analysis. 1982.（東南圖書公司）。

註：書籍可由括號中之書局買到。

4. Primitive Rings and Density Theorem

指導老師：呂溪木老師
四　甲：羅昭強

This paper is written for my graduation and thanks are due to all my professors in National Taiwan Normal University

Contents

1. Introduction

Throughout this paper, I introduce some of the basic building blocks of ring theory and provide some main theorems for primitive rings.

In the first part of my paper, we need some fundamental concepts of modules, and then, we go on discussing cur main topic ——Primitive Rings.

The aim is to introduce the famous theorem——Density Theorem, which was due to Jacobson and Chevalley. And I take it as the conclusion of my paper.

In this paper, my introduction is very concise. Everyone interested in this topic may read the references listed in the back of this paper.

2. Modules

Modules were considered in 1890's by Mibert and in the early part of this century by E. Neother. But the application of modules tc the study of the internal structure of rings was not realized until 1940's.

Def. 2-1

Let R be a ring. A non-empty set M is said to be a right R-module if M is an abelian group under an operation "+" such that for every r in R and m in M, there exists an element mr in M subject to:

(1) $(a+b)r = ar+br$

(2) $a(r+s) = ar+as$

(3) $a(r\cdot s) = (ar)s$

(4) $a1 = a$

for all $a,b \in M$; $r,s \in R$.

We will denote a right R-module M by M_R. A left R-module ${}_RM$ is defined symmetrically.

Example:

2-1 If R is a field, M_R is usually called a vector space over R.

2-2 If R=Z, the ring of integers, then M_R is just an abelian groups.

2-3 Any ring R is an R-module over itself.

2-4 Let F be the set of endomorphisms of the additive abelian group M written on the right. We obtain a ring (F, 0, 1, -, +, *) upon defining

$a0 = 0$

$a1 = a$

$$a(-f) = -(af)$$
$$a(f+g) = af+ag$$
$$a(f*g) = (af)g$$

for any $a \in M$ and $f, g \in F$. Moreover, we have the mapping

$$(a,f) \longrightarrow af$$

of $M \times F$ into M

Since f is a homomorphism,

$$(a+b)f = af+bf$$

hence, we have a right module M_F

Def 2-2:

Given two R-modules M and N, a function

$$f: M \longrightarrow N$$

is called a module homomorphism if

(1) $f(m+m') = f(m) + f(m')$

(2) $f(rm) = rf(m)$

for all $m, m' \in M$ and all $r \in R$.

By the definition, we say that the function is R-linear. As usual, if f is a module homomerphism of M into N, we define the kernel of f to be ker f= $\{x \in M \mid f(x)=0\}$. Clearly, it is a submodule of M and the image of f is also a submodule of N.

3. Primitive Rings

We begin with a basic concept in the structure theory of rings, this special rings we introduced play an important part in ring theory.

From now on, modules will be understood to be right modules.

Def. 3-1

A module M_R is called irreducible if it has exactly two submodules. Clearly, these submodules must be 0 and M

itself. The definition is meant to imply that $M \neq 0$.

Def. 3-2

An ideal P in the ring R is called (right) primitive if it is the largest ideal containrd in some maximal right ideal M.

Thm. 3-1

Let R be a ring with identity. P is a primitive ideal if and only if $P=R:M=\{r\in R \mid Rr\in M\}$ for some maximal right ideal M of R.

Pf: Only if part.

Suppose that P is primitive

There exists maximal right ideal M such that P is the largest ideal contained in M.

$\forall_r \in P \Rightarrow Rr \subset P \subset M$

$\Rightarrow r \in R:M$

$\Rightarrow P \subset R:M$

Suppose $P \subsetneq R:M$

$\exists\ r' \in R:M$ such that $r' \notin P$

$\Rightarrow P \subsetneq P+r'R+Rr'$

Since $r' \in R:M$

Then $Rr' \subset M$

$\Rightarrow r' \subset M$

$\Rightarrow r'R \subset M$

$\Rightarrow P \subsetneq P+r'R+Rr' \subset M$

Therefore P is not the largest ideal contained in M.

It is a contradiction.

If part.

Let $P=R:M+\{r\in R \mid Rr\subset M\}$

$\forall_r \in P \Rightarrow Rr \subset M$

$\Rightarrow r \in M$

$\Rightarrow P \subset M$

Claim: P is an ideal

Let $r_1, r_2 \in P$

We have $Rr_1 \subset M$ and $Rr_2 \subset M$

So $R(r_1-r_2)=Rr_1-Rr_2\subset M$

$\forall\ a\in R,\ r\in P$

$Rar\subset Rr\subset M$

$Rra\subset Ma\subset M$

Therefore, P is an ideal

claim: P is the largest ideal contained in M.

Suppose that Q is an ideal such that

$\forall\ s\subset Q,\ Rs\subset Q\subset M$

$\Rightarrow\ s\in P$

$\Rightarrow\ P=Q$

Hence, the results are following.

Def: 3-3

A ring is called a (right) primitive ring if O is a primitive ideal.

Clearly, we have some results about primitive rings.

(1) For any ideal P, R/P is a primitive ring if and only if P is a primitive ideal.

(2) A commutative ring is primitive if and only if it is a field.

In order to state another characterization of primitive rings, which is sometimes taken as definition, we need one more concept.

Def. 3-4

A module A_R is called faithful if for any $o\neq r\in R$, $Ar\neq O$.

Thm. 3-2

The ring R is primitive if and only if there exists a faithful irreducible module A_R.

Pf: Only if part.

Given O=R:M for some maximal right ideal M of R.

Try to exhibit a faithful irreducible module A_R.

Since (R/M, +) is an abelian group.

Make (R/M, +) into a right R-module

Let A=R/M

Make A into a right R-module as follows.

$\forall$ $a=r+M\in A$, $r'\in R$

Define $ar'=(r+M)r'$

$=rr'+M$

Claim: A is a right R-module

$r_1+M=r_2+M \Rightarrow r_1-r_2\in M$

$\Rightarrow (r_1-r_2)r'\in M$

$\Rightarrow r_1r'-r_2r'\in M$

$\Rightarrow r_1r'+M=r_2r'+M$

So it is well-defined.

Claim: A is an irreducible right R-module.

Since M is maximum right ideal of R.

Then R/M has no nontrivial right ideal.

Therefore R/M=A is irreducible

Claim: A_R is faithful

To do so, we just show that.

Ann $A=\{r\in R \mid Ar=0\}=0$

$\forall$ $r\in$Ann $A \Rightarrow A_r=0$

$\forall$ $x\in R$, $(x+M)r=xr+M$

$=M$

$\rightarrow R_r\subset M$

$\rightarrow r\in R:M$

$\rightarrow r=0$

If part

Given a faithful irreducible module A_R, that is, $A\neq 0$.

There exists an element a in A such that $a\neq 0$.

Define $\theta: R\rightarrow A$ to be

$\theta(r)=ar$

Claim: θ is module homomorphism

Let $r_1, r_2\in R$

$$\theta(r_1+r_2)=a(r_1+r_2)$$
$$=ar_1+ar_2$$
$$=\theta(r_1)+\theta(r_2)$$

Let $r, r' \in R$

$$\theta(rr')=a(rr')$$
$$=(ar)r'$$

Since $\theta(R)$ is sumbmodule of A and A_R is irreducible.

$\because$ $a \in \theta(R)$

$\therefore$ $\theta(R) \neq 0 \Rightarrow \theta(R) = A$

Thus $R/\ker\theta \simeq A$

A is irreducible

$(\Leftrightarrow)$ $\ker\theta$ is a maximum right ideal of R

Put $M=\ker\theta$

Claim $R:M=0$

$\forall\ r \in R:M \Rightarrow Rr \in M$

$\Rightarrow (R/M)r=M$

$\Rightarrow r \in \text{Ann } R/M$

$\Rightarrow r \in \text{Ann } A$

$r=0$

Hence R is primitive ring.

Thm.3-3

Any prime ring with a minimal right ideal is primitive.

Pf: Let M be the minimal right ideal of R

Then M_R is irreducible module

Claim: M_R is faithful

Let $r \in R$ such that $Mr=0$

$\Rightarrow MrR=0$

Since 0 is prime ideal

Then $rR=0$

$\Rightarrow r=0$

Therefore M_R is faithful

Hence R is primitive ring.

Thm.3-4

Any simple ring ($\neq 0$) is a primitive ring.

Pf: Since R is simple

Then 0 is a maximal ideal.

Clearly $R:0=0$

Hence R is primitive ring.

Def.3-5:

A ring R is said to be prime ring if and only if $aRb=0$; $a,b \in R$ implies that $a=0$ or $b=0$

Thm.3-5

A primitive ring is prime.

Pf: Let $P \neq 0$ be a right ideal of R

Suppose $Pa=0$

Since R is primitive ring, there exists a faithful irreducible R-module M_R

Since R is faithful on M

Then $MP \neq 0$

$\Rightarrow MP=M$

Then $Ma=(MP)a$

$=M(Pa)$

$=0$

It forces $a=0$

Hence R is prime ring.

Thm.3-6

The centre Z of a primitive ring R is an integral domain.

Pf: Let Z be the centre of prime ring R

And $a \in Z$; $b \neq 0$; $ab=0$

Then $0=Rab$

$=aRb$

The primeness of R and $b \neq 0$ implies $a=0$

The result follows

Example 3-1

The ring L of linear operators of a finite dimensional vector space over a division ring is simple, and hence, L is a primitive ring.

The concept of primitivity of a ring and of an ideal in a ring is not right-left symmetric.

For, there exist right primitive rings that are not left primitive in the sense that they have no irreducible left modules. The first examples of such rings were given by George Bergman.

4. Schur's Lemma

For an irreducible R-module A, the commuting ring turns out to be rather special. This is the content of an old and basic result known as Schur's Lemma.

Thm.4-1 (Schur's Lemma)

If A_R is an irreducible module, then its ring of endomorphisms $D=\mathrm{Hom}_R(A,A)$ is a division ring.

Pf: Let $0\neq d\in D$

Show that d^{-1} exists in D

$0\neq d \Rightarrow \mathrm{Im}\ d\neq 0$

$\Rightarrow \mathrm{Im}\ d=A$

$\Rightarrow$ d is onto

$0\neq d \Rightarrow \ker d\neq A$

$\Rightarrow \ker d=0$

$\Rightarrow$ d is 1-1

Since, d is 1-1, onto, then d^{-1} exists

Claim: such function d^{-1} is indeed R-linear

$\forall\ a_1,a_2\in A$

$d(d^{-1}(a_1)+d^{-1}(a_2))=d(d^{-1}a_1)+d(a^{-1}a_2)$

$=a_1+a_2$

$$\rightarrow d^{-1}(a_1+a_2)=d^{-1}(a_1)+d^{-1}(a_2)$$

$\forall\ a\varepsilon A,\ r\varepsilon R$

$$d(a^{-1}(a)r)=d(d^{-1}(a))r$$
$$=ar$$
$$\rightarrow d^{-1}(ar)=d^{-1}(a)r$$

From above, we know that $d^{-1}\varepsilon D$

Since, every non-zero element in D has inverse, so D is a division ring.

Def. 4-1

Given rings A, B and abelian group M.

M is an (A,B) bimodule if and only if

(1) ${}_AM$ is left module

(2) M_B is right module

(3) a(mb)=(am)b

for all $a\varepsilon A$, $m\varepsilon M$, $b\varepsilon B$.

Now, we give an example of bimodule and do some preparation for our next section.

Given a right primitive ring R, say with faithful irreducible module M

Put $D=Hom_R(M,M)$; endomorphism ring of M_R

By Schur's Lemma, D is division ring

We can considered M as left D-module

${}_DM$: $d\cdot m \longrightarrow d(m)$

That is, ${}_DM$ is a left vector space

Then ${}_DM_R$ is a (D,R) bimodule

$$(d\cdot m)r=(d(m))r$$
$$=d(mr)$$
$$=d\cdot(mr)$$

We have another ring as well as D

Put $E=Hom_D(M,M)$; the ring of all linear operators of ${}_DM$

E is somewhat a "nice" ring.

We try to embede R into E.

Define: $\theta: R \longrightarrow E$ such that

$\theta(r): M \longrightarrow M$

given $m \cdot \theta(r) = mr$

We will show that $\theta(r)$ is a D-linear

$\forall\ m, m' \in M,\ d \in D$

$$(m+m')\theta(r) = (m+m')r$$
$$= mr + m'r$$
$$= m\theta(r) + m'\theta(r)$$
$$(dm)\theta(r) = (d(m))\theta(r)$$
$$= (d(m))r$$
$$= d(mr)$$
$$= d \cdot (m\theta(r))$$

We will show that θ is ring-homomorphism

$\forall\ m \in M;\ r_1, r_2 \in R$

$$m\theta(r_1 + r_2) = m(r_1 + r_2)$$
$$= mr_1 + mr_2$$
$$= m\theta(r_1) + m\theta(r_2)$$
$$= m(\theta(r_1) + \theta(r_2))$$
$$\theta(r_1 + r_2) = \theta(r_1) + (r_2)$$
$$m\theta(r_1 r_2) = m(r_1 r_2)$$
$$= (mr_1)r_2$$
$$= (m\theta(r_1))r_2$$
$$= (m\theta(r_1))\theta(r_2)$$
$$= m(\theta(r_1)\theta(r_2))$$
$$\Rightarrow \theta(r_1 r_2) = \theta(r_1)\theta(r_2)$$

We also show that θ is 1-1

$$\forall\ r \in \ker\theta \Rightarrow \theta(r) = 0 \in E$$
$$\Rightarrow m\theta(r) = 0 \quad \forall\ m \in M$$
$$\Rightarrow mr = 0 \quad \forall\ m \in M$$
$$\Rightarrow Mr = 0$$
$$\Rightarrow r = 0 \quad (\because M \text{ is faithful})$$

So, if R is a primitive ring with faithful irreducible module M_R, then R may be considered as a subring of E, where

$E=\mathrm{Hom}_D(M,M)$; $D=\mathrm{Hom}_D(M,M)$

${}_DM_R$ is bimodule.

Thus R is embedded into the linear transformation ring E. But how "densely" that R is embedded into E? Jacobson and Chevalley proved that it was very "densely" embedded!

5. Density Theorem

If V is a vector space over a division ring D, then a set S of linear transformations in V is called dense if for any given finite set of linearly independent vectors x_1, x_2, ..., x_n and corresponding vectors y_1, y_2, ..., y_n, there exists an $\ell \in S$ such that $x_i\ell=y_i$, $1\leq i\leq n$.

If V is finite dimensional, then we can take the x_i to be a base and $y_i=x_ia$ for any given linear transformation a. Then we have an $\ell \in S$ such that $x_i\ell=x_ia$, $1\leq i\leq n$, from which it follows that $\ell=a$. Hence the only dense set of linear transformations in a finite dimenional vector space is the complete set of linear transformations.

In this section we shall prove a basic theorem that permits the identificution of any primitive ring with a dense ring of linear transformations.

First, we give the definition and equivalence relation to the word "dense".

Def.5-1

Let R be primitive with $R\subset E$, where

$E=\mathrm{Hom}_D(M,M)$; $D=\mathrm{Hom}_R(M,M)$

${}_DM_R$ is bimodul; M_R is faithful irreducible module of R

(1) R is said to act densely on E if G is a finite generated submodule of ${}_DM$ and $e\in E$, then $\exists\ r\in R$ such that

$$G(e-r)=0.$$

(Or equivalently)

(2) For each positive integer n, each linearly independent subset $\{m_1,m_2,\ldots,m_n\}$ of ${}_DM$ and each subset $\{w_1,w_2,\ldots w_n\}$ of ${}_DM$, there exists $r\varepsilon R$ such that

$$m_i r=w_i.$$

For the above statements, we give the proof as follows:

Pf: 1 $\Rightarrow$ 2

Let $G=\{m_1,m_2,\ldots,m_n\}$

Since G is linearly independent

There exists an element $e\varepsilon E$ such that

$$m_i e=w_i \quad i=1,2,\ldots n$$

By hypothesis, there exists $r\varepsilon R$ such that

$$G(e-r)=0$$

For all i, $m_i\varepsilon G \Rightarrow m_i(e-r)=0$

$$\Rightarrow m_i e-m_i r=0$$

$$\Rightarrow w_i-m_i r=0$$

$$\Rightarrow w_i=m_i r.$$

2 $\Rightarrow$ 1

Suppose $G=\{m_1,m_2,\ldots,m_n\}$ be finitely generated

For any $e\varepsilon E$, there exists a set $\{w_1,w_2,\ldots w_n\}$ such that

$$m_i e=w_i \quad i=1,2,\ldots,n$$

By hypothesis, there exists $r\varepsilon R$ such that

$$m_i r=w_i \quad i=1,2,\ldots,n$$

Therefore, we have

$$m_i e=m_i r \quad \forall_i$$

$$m_i(e-r)=0 \quad \forall_i$$

$$G(e-r)=0$$

To prove Density Theorem, we need the following pair of definitions.

Def.5-2

For any subset S of M,

define

$$S^r=\{x\in R \mid Sx=0\}$$

For any subset T of R

define

$$T^\ell=\{x\in M \mid xT=0\}$$

Obviously, S^r is a left ideal of R and T^ℓ is a submodule of M.

Thm.5-1 (Density Theorem)

Let R be a primitive ring with faithful irreducible module M_R. May assume $R\subset E$ where $E=Hom_D(M,M)$; $D=Hom_R(M,M)$

Then, for each $e\in E$ and every finitely generated such module G of ${}_DM$, there exists $r\in R$ such that

$$G(e-r)=0$$

Pf: We take induction on the number n of generators of G to show that

(1) $G(e-r)=0$

(2) $G^{r\ell}=G$

(a) Suppose n=0

Then G=0 and $G(e-r)=0$

$$(G^r)^\ell=R^\ell$$
$$=0$$
$$=G$$

(b) Assume that $n\neq 0$ and that (1), (2) hold for all submodules of ${}_DM$ with the number of generators less than n.

Let $G=[a_1,a_2,\ldots,a_n]$ be a submodule of ${}_DM$. May assume that $a_1,a_2,\ldots,a_n$ are linearly independent over D.

$$G=\sum_{i=1}^{n} Da_i$$

$$=G'+Da_n$$

$$a_n \triangleq a$$

$$=G'+Da$$

$a_n \notin G' = \langle a_1, a_2, \ldots, \hat{a}_n \rangle$

Let $e \in E$, then by induction hypothesis E $s \in R$ such that

(1) $G'(e-s)=0$

(2) $G'^{r\ell}=G'$

Try to exhibit $r \in R$ such that

$G(e-r)=0$

$\Leftrightarrow$ $(G'+Da)(e-r)=0$

$\Leftrightarrow$ $(G'+Da)(e-s+s-r)=0$

$\Leftrightarrow$ $(Da)(e-s)=(G'+D_a)(s-r)$

$\Leftrightarrow$ $(Da)(e-s)=G'(s-r)+(Da)(s-r)$

It suffices to exhibit $r \in R$ such that

$G'(s-r)=0$ and $a(s-r)=a(e-s)$

Let $b=s-r$

That is, $G'b=0$ and $ab=a(e-s)$

It suffices to exhibit $b \in R$ such that

$G'b=0$ and $ab=a(e-s)$

This is possible if ${}_aG'^{r}=M$

Since $a(e-s) \in M$

If $aG'^{r}=M$

Then, there exists $b \in G'^{r}$ such that

$G'b=0$ and $ab=a(e-s)$

Now, we claim that $aG'^{r}=M$

Suppose not, $aG'^{r}=0$ because M_R is irreducible and aG'^{r} is an R-module of M

We will show that ${}_aG'^{r}$ is submodule of M.

Let $x_1, x_2 \in G'^{r}$

Then $G'x_1=0$, $G'x_2=0$

$\Rightarrow G'(x_1+x_2)=0$

$\Rightarrow x_1+x_2 \in G'^{r}$

So $a(x_1+x_2)=ax_1+ax_2 \in aG$

Let $ax \in aG'^{r}$, $y \in R$

Then $(ax)y = a(xy)$

$\forall\ g \in G$, $g(xy) = (gx)y$

$= 0\ y$

$= 0$

$\Rightarrow xy \in G'^{r}$

$\Rightarrow aG'^{r}$ is R-module

But $G'^{r} = 0$

$\Rightarrow a \in G'^{rl} = G'$

$\Rightarrow a \in G'^{lr}$ $\nrightarrow\nleftarrow$

That is, $aG' = M$

There (1) holds.

Remains to show that $G^{rl} = G$

Clearly $G^{rl} \supset G$

Since $\forall g \in G$, $\forall x \in G^{r}$

We have $gx = 0$ which implies

$G^{rl} \supset G$.

On the other hand.

If $b \in G^{rl}$, then

$b \in (G'+Da)^{rl} = [(G'+Da)^{r}]^{l}$

For $x \in (G'+Da)^{r}$

$\Rightarrow (G'+Da) = 0$

$\Rightarrow G'x + Dax = 0$

$\Rightarrow G'x = 0$ and $Dax = 0$

$\Rightarrow x \in G'^{r} \cap (Da)^{r}$

Therefore, we have

$b \in G^{rl} \Rightarrow b \in [(G'+Da)^{r}]^{l}$

$\Rightarrow b \in (G'^{r}\ Da^{r})^{l}$

$\Rightarrow b \in (G'^{r}\ \{a\}^{r})^{l}$

Thus $b_x = 0$ whenever $G'x = 0$ and $ax = 0$

Consider submodule aG'^{r} of M_R

Since $aG'^{r} = 0$ or $aG'^{r} = M$

(i) If $aG'^{r} = 0$, then $bG'^{r} = 0$

$\rightarrow b \in G'^{rl} = G' \subset G$

$\rightarrow b \in G$

(ii)If $aG'^{r}=M$, then

define:

$d: aG^{lr} \to bG'^{r}$ such that

$ax \xrightarrow{d} bx$

We will show that d is well-defined.

For $ax_1=ax_2$

$\Rightarrow a(x_1-x_2)=0$

$\Rightarrow b(x_1-x_2)=0$

$\Rightarrow bx_1=bx_2$

$\Rightarrow d(ax_1)=d(ax_2)$

We will show that $d \in \mathrm{Hom}_R(M,M)=D$

$\forall\ x \in G'^{r}$

$bx=d(ax) \in {}_DM_R$

$=(da)x$

$\Rightarrow (b-da)x=0$

$\Rightarrow (b-da)\ G'^{rl}=G'$

$\Rightarrow b \in G'+da$

$\Rightarrow b \in G'+Da=G$

$\Rightarrow G^{rl} \subset G$

Therefore (2) holds.

The density theorem allows us to draw many conclusions about primitive rings and to relate them to matrix rings. We introduce a theorem to show this.

Thm. 5-2

Let R be a primitive ring. Then for some division ring D, either R is isemorphic to D_n, the ring of all nxn matrices over D, or given any integer m, there exists a subring S_m of R which maps homomorphically onto D_m.

Pf: Since R is a primitive ring.

By the density theorem, R is a dense ring of linear transformations on a vector space V over a division ring D.

If V is finite dimensional over D.

Then, the density of R on V tells us that R is isemorphic to the ring of all D-linear transformations on V (In the preliminaries of section five)

That is, $R \cong D$ where n=dim V.

If V is not finite dimensional over D.

Then, given any positive integer m, we can find m linearly independent elements $v_1, v_2, \ldots, v_m \in V$.

Let V_m be the subspace of V generated by $v_1, v_2, \ldots v_m$.

Let $S_m = \{x \in R \mid V_m x \subset V_m\}$

The density theorem says that any D-linear transform can be induced by an element of R.

Thus, S_m is precisely the set of all elements of R which induce linear transformations of V_m.

Therefore, we have a surjective homomorphism

$$S_m \longrightarrow D_m$$

If $W_m = \{x \in S_m \mid V_m x = 0\}$

Then, we have $S_m / W_m \cong D_m$

Thus, the result follows.

6. References

(1) Burton. D.M.
A First Course In Rings And Ideals

(2) Gray. M.
A Radical Approach To Algebra

(3) Herstein, I.N.
Noncommutative Rings

(4) Jacobson. N.
Basic Algebra

(5) Lambek. J.
Lectures On Rings And Modules.

(6) Notes from professor Hsi-Muh Leu.

發 行 人：顏 啟 麟
出 版 者：國立臺灣師範大學數學學會
主　　編：許 志 農
封　　面：吳 基 國
印 刷 者：九章文具印刷品有限公司
台北市虎林街 252 巷 81 弄 33 號 2 樓
電　　話：(02)704-5243
出版日期：中華民國七十四年六月廿五日
師大訓課刊登第 136 號