統計数理研究所彙報
第31巻 第1号 (1983)

パーソナル・コンピュータのための 物理乱数発生器

統計数理研究所 仁 木 直 人

(1983年5月 受付)

1. はじめに

漸化式を用いて発生する算術乱数を大量に使用するとき，不都合な性質が現われることがある．特に，乗数と法の組合せが不適当であると，乗算合同法で得られる乱数に顕著な「結晶構造」が現われて使用に耐えないこと (Marsaglia [7], Knuth [6]) は良く知られるようになった．また，乗算合同法に代わるべきものとして導入された m 系列による発生法（発生法自体は1950年代から良く使われているが，実用的になったのは Arvillias and Maritsas [1] からで，伏見・手塚 [3] にその改良がある）は，後に見るように乱数列の統計的性質があまり良くないようである．

これに対して，物理現象を元に生成した乱数列は，その発生法が適切であれば，全く理想的な性質を持つことが期待でき，発生速度も大型計算機用として充分なものが製作できる（石田 [4]，仁木 [9]）．

パーソナル・コンピュータでは，その主要な言語である BASIC に組込まれた関数 RND を一様乱数源として用いることが圧倒的に多い．この RND 関数も乗算合同法を基礎とした算術乱数の一種である．最近は，高度なゲームに用いられるせいか，その統計的性質にも注意が向けられ，相当複雑な発生方式が採用されている．しかし，複雑にした（かつ遅くなった）割には，より単純な方式に比べてあまり性質が向上したとは言えず，シミュレーションなど，乱数の性質が重要となる問題に使うことはできない．

そこで，できるだけ簡略化した構造を持つ物理乱数発生器を試作し，そこから得られる乱数列の性質を調べ，他の発生法との比較を行なうことを考えた．

パーソナル・コンピュータ用という点に留意して，①特殊な部品や器具を必要としない，②容易かつ安価で製作できる，③調整個所が少なく容易で，また再調整の必要がほとんどない，④小型でコンピュータ本体内に収納できる，⑤10進または16進乱数の選択ができる，⑥発生速度を1,000桁/秒以上とする，⑦乱数の精度は，相対誤差 10^{-3} 以下とする，などの仕様を定めて回路設計を行ない，テスト回路を付加した発生器を試作した．

発生原理は新しいものではないが，乱数源，2値パルス化，乱数値取出し部の各回路には，性能を落とすことなく簡素化を図る工夫を行なっている．その詳細を第2章で述べる．

この試作器を用いて発生される，16進乱数およびその元となるランダム・パルスの分布に関し，第3章に計測結果を報告し他の乱数発生法との比較を行なっている．

2. 物理乱数発生器の原理と構造

2.1. 乱 数 源

乱数源となるのは，ツェナ・ダイオード（定電圧ダイオード）が適当な逆バイアス条件下で

発生する雑音で，かなり広い周波数帯域において白色雑音と見なせるものである．この乱数源を用いる利点は多く，①高い周波数帯域を利用することにより高速の乱数発生が可能なこと，②簡単な回路で小さく組上がること，③特殊な部品・材料を必要としないこと，④逆バイアス条件の調整は，最適点から多少ずれても動作に支障がないので，比較的容易であり，また再調整の必要がほとんどない，など特別な用途（例えば超高速発生）を除けば最適なものと言えよう．

一部の人が危惧を抱く長期安定性に関しては，統計数理研究所の大型コンピュータに接続されている乱数発生装置（回路構成は異なるが，乱数源としてツェナ・ダイオードを用いている）が，ほとんど何の調整もなく，15年近く良好な状態を保っていることを挙げておく．

他の良く使われる乱数源は，パーソナル・コンピュータ用として不適当と思われる．放射性物質を使用するものは，発生速度が非常に遅い上，放射線の検出器という高価で大きな器具を必要とする．サイラトロンや他のガス封入放電管の使用も可能であるが，別の電源を必要とし，回路も大型とならざるをえない．高周波ダイオードからの雑音は，周波数帯域が高域まで広がり乱数源として好ましいものであるが，雑音レベルが非常に小さいので，その増巾回路は設計・製作ともに専門家にとっても難しく，結局大がかりにならざるを得ない．

試作器では，回路を特に簡単にするため，ツェナ・ダイオード（RD9A）を直接トランジスタ（2SC717）のベースに直列接続し，トランジスタの動作点をエミッタにつけた可変抵抗器で調整する構成とした（図1参照）．トランジスタのコレクタから，簡単なハイパス・フィルタを通して，100～300 mV の雑音が得られる．可変抵抗器の調整は，この雑音波形をオシロスコープで見ながら，最も出力が大きくなるように行なう．広い範囲でほぼ最大の出力が得られるから，調整は簡単である．ただし，発振していないことを波形から確認する必要がある．発振防止には，雑音源回路を小さくまとめ，回路近くで 12 V の電源にバイパス・コンデンサを入れる．調整しても充分な雑音出力が得られない場合は，ツェナ・ダイオードを取換えてみる．

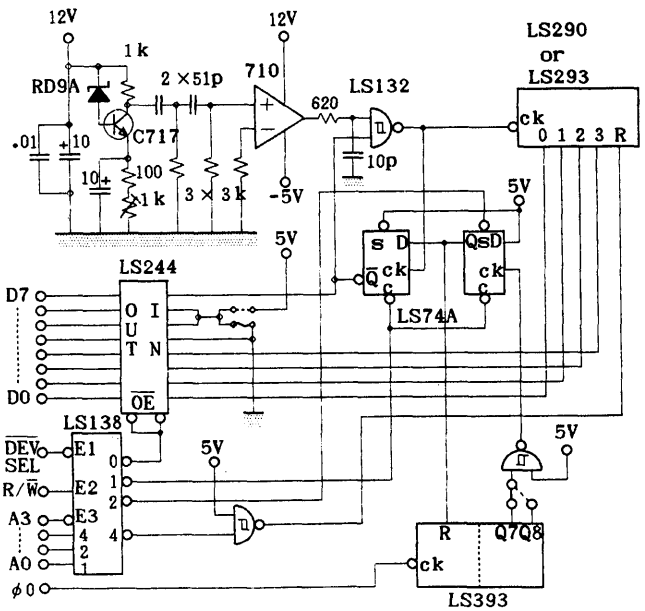

図1　簡単な10進/16進物理乱数発生器（8000～4000個/秒）

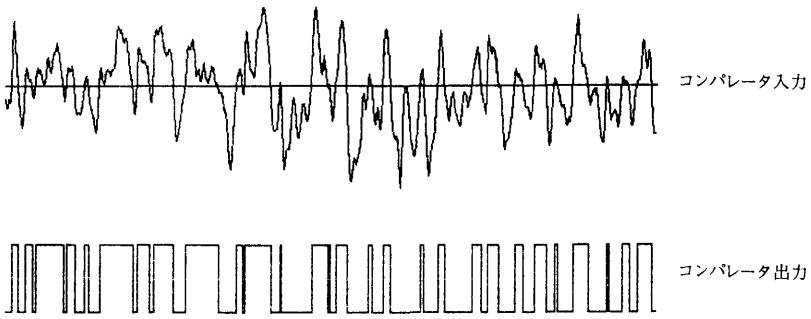

図2　ツェナ・ダイオード雑音の2値化

数本の内には，雑音を比較的強く出すものがあると思う．

このようにして得られた雑音をコンパレータ (710) により2値化する．ここでは，雑音電圧の正負により，1または0の論理レベルを持つパルス列としている (図2)．もし，雑音源を2組用意できれば，その両出力をコンパレータで比較する方式とし，信頼性の向上を図ることができよう．

なお，次段のカウンタ部の動作を確実にするため，積分回路とシュミット・ゲート (74LS132) を用いて，巾の極端に狭いパルスの除去を行なっている (図3)．

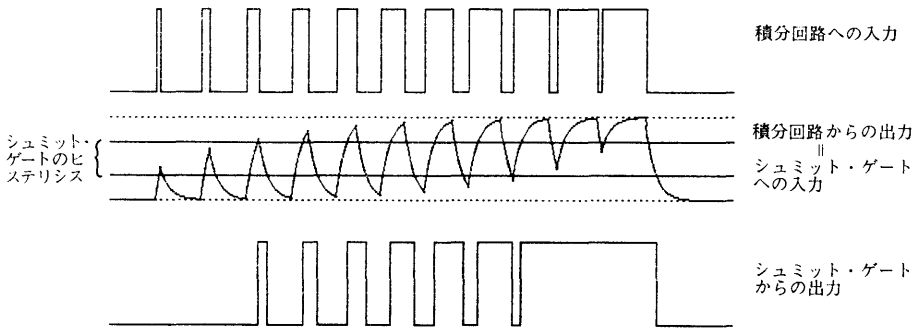

図3　積分回路とシュミット・ゲートによる極端に巾の狭いパルスの除去

この結果，乱数源として数百万パルス/秒の (ON/OFF 可能な) ランダム・パルスが得られる．このパルスの一定時間当たりの頻度分布については，3.1節で述べる．

2.2.　10進または16進乱数の生成

乱数源より得られるランダム・パルスを一定時間計数し，その計数値の下一桁を乱数の値として採用することにする．計数を何進法で行なうかによって，2進，10進あるいは16進などの乱数が得られよう (仁木 [9], [10])．通常はこれらを数個組合せて必要な範囲の値を持つ乱数を合成することになる．

高速性と回路の簡単さからは，2進乱数を発生するのが最も有利であるが，2進1桁の乱数をそのまま使うことは稀で，数桁連結する必要がある．大型計算機用では，独立な2進乱数発

生回路を多数並列に配置することにより，高速性を維持している（石田ら [5]）．しかし，パーソナル・コンピュータに用いるには，あまり速度の面を考える必要はなく，1秒間に数千個発生できれば充分であろう．そこで，試作器では，組合せて使うばかりではなく，そのままでも使われることの多い10進乱数および16進乱数を直接発生することとした．

10進あるいは16進の計数には，74LS290（10進）または74LS293（16進）を用いている．この両者は，入出力や電源などのピン配置がほぼ同等であるから，10進と16進の切換えはICを交換することで行なう．上位桁は乱数値としては不要であるので，一桁分だけあれば良い．

計数時間は，組込むパーソナル・コンピュータのクロックが利用できれば，その整数倍の適当な値を採用する．計数時間が不足すれば得られる乱数の一様性に問題が生じるし，長すぎれば発生速度が低下する．

試作器では，1.023 MHz のクロックを2進カウンタ（74LS393）で分周して，125μ秒または250μ秒の計数時間を得ている．125μ秒（8,000桁/秒）の計数時間では，得られる乱数の分布と離散一様分布との相対誤差は10^{-3}以下，250μ秒（4,000桁/秒）では10^{-6}以下と計算でき，パーソナル・コンピュータ用には充分であろう．なお，精度等については第3章で詳しく述べる．

カウンタを停止させる際には，カウンタが安定状態にある時に行ない，回路や素子の非対称性が乱数値に影響することを防いでいる（仁木 [9]，[10] 参照）．具体的には，カウンタへの入力パルスの立上がりエッジと同期して停止を行ない，立下がりエッジで計数値が変化する過程が充分安定した時点を捕える．事前に巾の極端に狭いパルスを取除いた理由は，安定するまでの時間を確保する必要からである．

なお，カウンタは帰零できるようにすべきである．この配慮は，通常の使用のためではなく，乱数の性質の検定を行なう際に，一様性と独立性を明確に分離するために必要となる（仁木 [9]）．

2.3. 計算機との接続

組込むパーソナル・コンピュータによって，乱数発生器の接続方法が異るのは当然である．共通することは，次のような過程の存在である．①CPU 側からの発生開始指令⟶②発生器からの発生終了報告⟶③データ（乱数）の読み込み．この他，特別な制御が必要になることもあろう．試作器では，乱数の検定に便利なように，①②③に対応する機能の他に，カウンタの帰零を行なう指令などを受付けるようになっている．

試作器は APPLE II の周辺機器用スロット（Peripheral Connectors）に直接入れられるようなインタフェース構成をとり，次の四種の指令が使用できる．

① 発生開始（code=1）
② データ読み込み（code=0）
③ 発生の緊急停止（code=2）
④ カウンタの帰零（code=4）

各指令は

$(C080)_{16}$+(スロット番号)×$(10)_{16}$+code

番地を read することにより実行される．データ読み込み②により，発生器の状態および現在のカウンタ値がAレジスタに読み込まれる．Aレジスタの内容は次のような意味を持つ．

2^7……0 ならば発生終了，1 ならば発生中
2^6, 2^5……常に 0（または常に 1）
2^4……常に 0

2^3～2^0･･･乱数値（ただし，2^7=0 のときのみ有効．10 進数は BCD コードである．）

2^7 ビットは，発生開始①により 1 に設定され，125（または 250）μ 秒経過後の最初の立上がりエッジで 0 に戻る．時間経過前に緊急停止指令③が出されたときは，直ちに 0 に設定される．2^6 および 2^5 ビットは通常 0 としておくが，10 進数を ASCII コードの“文字”として読みたいときは 1 とすることも便利であろう．これは回路上でジャンパ切換を行なうことにより可能である．

他機種のパーソナル・コンピュータへの接続も，試作器のインタフェース回路を手直しすることにより，各種の接続方法で実現できよう．例えば，ROM（読出し専用メモリ）のソケットを利用する方法などである．回路全体を一枚の小さなボード上に製作することは容易であるから，なるべくならパーソナル・コンピュータの本体に内蔵可能にすることである．

3. 乱数列の統計的性質

3.1. ランダム・パルス

ランダム・パルスの 1 計数時間 125μ 秒当たりの発生頻度は，使用するツェナ・ダイオード他の素子の特性にもよるが，平均 430～470，分散 100～120 の正規型分布に従う．あるいは，平均のずれたポアソン分布といった方が良いかも知れない．因みに，歪度 $\sqrt{b_1}$，尖度 b_2 は，20 回の測定結果（各回ともサンプル・サイズは 2^{20}）では，

$$|\beta_1|<0.005, \quad |\beta_2-3|<0.008$$

であった．平均カウント数は温度係数（約 −1/度）を持ち，周囲温度が上昇すると減少する．分散の大きさと周囲温度の関係は，20 回の実測からははっきりした傾向が見られない．

本来，単位時間当たりのカウント数は，ランダム・パルスが完全白色に充分近ければ，通常のポアソン分布（平均=分散）に従うと考えられる．平均≒分散となっていない主な理由は，使用する雑音の周波数帯域をかなり狭く設定してあることによる．すなわち，低周波成分を取除くことによってパルス数の平均・分散を増して高速化を図り，同時に論理回路の安定動作のために高周波成分を除去したためである．

図 4 に 125μ 秒間に生起したパルス数の分布を示す．計測回数は 2^{20}=1,048,576 回で，カウント数の最小は 415，最大は 511 であり，

平均	461.4	分散	109.9
$\sqrt{b_1}$	0.002	b_2	2.998

であった．正規性を仮定するとき，歪度および尖度の標準偏差はそれぞれ 0.002，0.005 と計算されるから，上記の $\sqrt{b_1}$，b_2 は妥当な値といえる．また，この分布を正規分布から得られたヒストグラムと考えて，その χ^2 適合度検定を行なうと χ^2=97.6（自由度 83）となり，良く適合しているといえよう．

3.2. 理論的精度

1 計数時間内に生起するパルス数の分散（あるいは標準偏差）の値が，生成される乱数の精度を決定する．当然のことながら，分散（あるいは標準偏差）が大きいほど離散一様分布からのずれは小さくなる．

ツェナ・ダイオードを何個か試してみて雑音レベルの大きいものを選び，ハイパス・フィルタの CR 値を適当に動かせば，125μ 秒間に生起するランダム・パルス数の分散を 100 以上とすることは，比較的容易である．

ここで，標準偏差 σ の正規分布に従う確率変数を Z とし，Z を n で除した非負の剰余

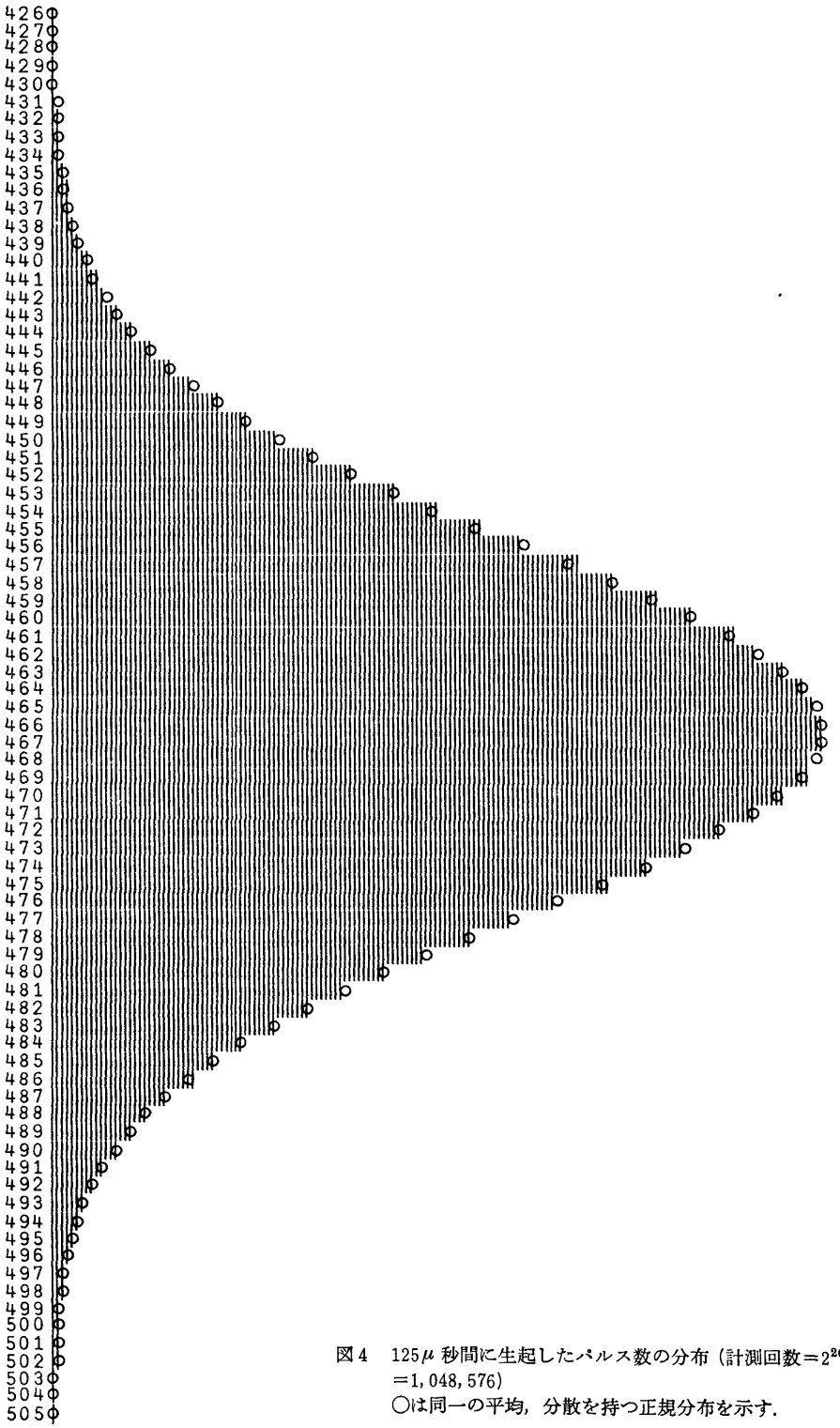

図4 125μ 秒間に生起したパルス数の分布（計測回数$=2^{20}$ $=1,048,576$）
○は同一の平均，分散を持つ正規分布を示す．

$$Y = Z - mn \quad (0 \leqq Y < n\ ;\ m \text{ は整数})$$

を考えると，Y は $[0,\ n)$ 上の一様分布に近似した確率密度

$$f(y) = \frac{1}{n}\left\{1 + 2\sum_{r=1}^{\infty} \exp\left(-\frac{\pi^2\sigma^2}{2n^2} r^2\right) \cos\ (y \text{ の一次式})\right\}$$

を持つ分布に従う (Niki [8])．$f(y)$ と $[0,\ n)$ 上の一様分布との相対誤差の上限は，

$$|nf(y) - 1| \leqq 2\sum_{r=1}^{\infty} \exp\left(-\frac{2\pi^2\sigma^2}{n^2} r^2\right)$$

によって与えられるが，右辺の和の $r \geqq 2$ の部分は，σ/n が 1/2 以上ならば，第1項の $1/10^6$ 以下であるから，実質上は第1項のみで評価すれば充分である．

標準偏差 σ のカウント数を持つランダム・パルスを用いて，n 進乱数を 2.2 節の方法により生成すると，その分布は Y の整数部 $\|Y\|$ ($\|\ \|$ はガウスの記号) の分布で良く近似されよう．よって，生成される n 進乱数 X の従う分布と $[0,\ n-1]$ 上の離散一様分布との相対誤差は，

$$\sup|n \Pr\{X = x\} - 1| \fallingdotseq 2\exp\left(-\frac{2\pi^2\sigma^2}{n^2}\right) \quad (x = 0,\ 1,\ \cdots,\ n-1)$$

で実質上評価される．

最悪のケースとして，$\sigma = 10$ と取り，16 進乱数を発生する場合 ($n = 16$) を考えると，この相対誤差は約 9×10^{-4} すなわち 0.1% 程度である．偶然による変動が 0.1% 程度に小さくなり，この相対誤差が問題となってくるのは，乱数を 10^6 個以上使用する場合である．従って，パーソナル・コンピュータ用としては充分であろう．もし，乱数を磁気テープなどに保存して大量に何回も使用するときなど，より高い精度を必要とするならば，発生速度を 4,000 個/秒 (250 μ 秒/個) とすれば良い．この場合の相対誤差は，10^{-6} 以下である．

なお，10 進乱数の場合の相対誤差は，それぞれ

125μ 秒/個のとき　10^{-8}
250μ 秒/個のとき　10^{-16}

程度と計算できる．

3.3. 実験と比較

理論的考察による乱数の精度は，3.2 節で述べたように充分満足できるものである．ここではその確認のための実験結果を報告し，数種の算術乱数との比較をも試みる．

物理乱数は，本来独立と見られる事象を利用して発生されるものであるから，「独立性」についての憂慮はほとんど必要としない．事実，1,000,000 個程度の乱数を用いた種々の独立性の検定では，全く独立性の崩れを見出すことができなかった．むしろ，回路の動作の非対称性に基く「一様性」の崩れの方が心配である．

大量の乱数を用いた実験では，3つの観点からデータを収集した．回路構成の不備があればその影響が最も出やすいと思われる最下位桁に関しては，「偶数・奇数の出現率」および「偶数・奇数出現の独立性」の2項目について計測した．また，16進乱数の「一様性」に関する計測も同時に行なった．「一様性」については，比較のため，数種の算術乱数を用いて同様の計測を行なった．

実験は，125μ 秒ごとに 16 進乱数1個を発生する条件で行なった．この条件設定を選んだのは，

① この試作器に関しては，最悪の条件であること，

② 比較する算術乱数に関しては，10進乱数を発生するのに比べれば，性質の良い乱数が得られやすい条件であること．

によるもので，物理乱数側に厳しい設定である．なお，乱数の性質を調べる際には，必ずカウンタを帰零してから各回の乱数発生を行なっている．これは，前述のように，「一様性」と「独立性」を明確に分離するためである．

3.3.1 実験の方法

実験は次のような手続きに従い，約1億2千万個の16進1桁乱数を用いて行なった．

① 乱数を初期値として1個発生する．

② 4,096個の乱数を発生させる．

③ そのうちに含まれる偶数の個数 E を求め，

$$\chi_1^2 = \frac{2}{2,048}(E-2,048)^2$$

を計算する．

④ ひとつ前に発生された乱数と現在の乱数の偶奇性に基き，それらを「偶偶」「偶奇」「奇偶」「奇奇」の4種に分け，各種別に属すペアの数（それぞれ P_0, P_1, P_2, P_3 とする）を求め，

$$\chi_3^2 = \frac{1}{1,024}\sum_{i=0}^{3}(P_i-1,024)^2$$

を計算する．

⑤ i なる16進値を取った個数 H_i を求め，

$$\chi_{15}^2 = \frac{1}{256}\sum_{i=0}^{15}(H_i-256)^2$$

を計算する．

⑥ (②～⑤) を30,000回くり返して，各 χ^2 値の度数分布を求める．同時に全体での偶数の個数（すなわち E の和）も求める．

4,096個の乱数発生に要する時間は約0.5秒と短く，この間の周囲温度・湿度および各素子の温度などの環境要因はほぼ一定と考えられる．それゆえ，環境要因により「一見性質が良く見える」現象は排除されよう．

上記のような手続きとした理由には，

① もし乱数としての性質にどこか不都合があったとしても，それは極く僅かな差異として検出される程度であろうから，その発見には大量の乱数と検出力の高い方式が必要となる．ここでは，この試作器と同程度の乱数発生器が要求されるであろう発生量の上限を100万個程度と考え，その約100倍の乱数を使用した．また，χ^2 値の分布形を見ることは，これまでの経験では，有意義な情報をもたらしてくれることが多かった．

② パーソナル・コンピュータを用いるため，内部記憶容量および計算速度の面で制約がある．それゆえ，χ^2 値の分布に基く検定を行なう場合も，例えば尤度比検定などは採用できず，ひとまず度数分布表の形に集約しておく形式とした．

の2つの側面がある．因みに，実験に要した時間は約11時間で，乱数発生に純粋に要する時間約4時間半の2倍以上になる．

3.3.2 偶数および奇数の出現率

実験に使われた 4,096×30,000=122,880,000 個の乱数のうち，偶数は 61,444,920 回現われた．偶数の出現が独立で確率 1/2 であるとする仮説の下では，出現度数は，平均 61,440,000，標準偏差 5,543 を持つ二項分布に従う．よって，この実験での実現値は，仮説が正しいとしたときに，充分に起こり得る値である．

次に，短いスパンでの偶数出現率を見ることにする．乱数 4,096 個ごとに求めた χ_1^2 (3.3.1 ③) の値は，上記の仮説の下で，自由度 1 の χ^2 分布に近似した分布に従う．一方，χ_1^2 の値の分布は，パーソナル・コンピュータの能力上の制約から，0.5 きざみでクラス分けした度数分布表の形で与えられる（表 1）．単純に「この度数分布表が自由度 1 の χ^2 分布から得られた」との仮説を置いて，検定を行なおうとするのは正しくない．その理由は，4,096 個中の偶数の個数は当然整数値しかとらないから，χ_1^2 のとり得る値も離散的であるからで，特に自由度が小さい場合にはその影響が大きい．ここでもその影響を考慮しなければ 5% 程度の誤差が入る可能性がある．表 1 には，各クラスについて，実現度数と離散性を考慮した期待度数および自由度 1 の χ^2 分布に基く期待度数を掲げておいた．

ヒストグラム（図 5）を見れば，実験から得られた χ_1^2 の度数分布は，「偶数の出現が独立で

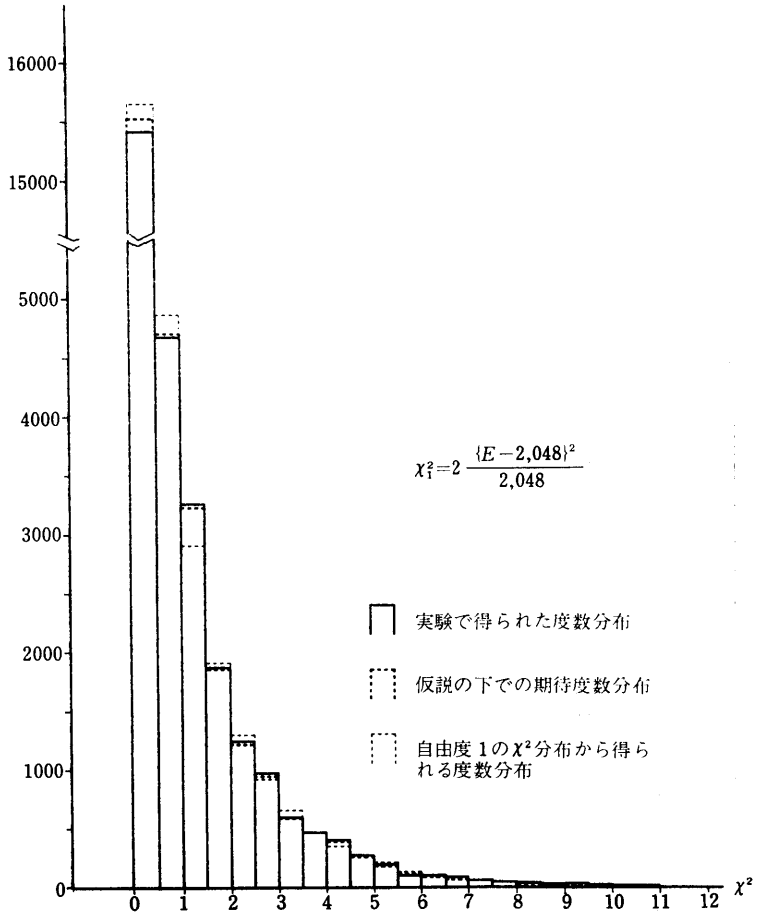

図 5　乱数 4,096 個ごとに求めた χ_1^2 値の度数分布

表 1. 乱数列の短い構成単位での偶数出現率の分布

発生した乱数列の連続する4,096個ごとに，偶数の出現回数 E から

$$\chi_1^2 = 2\times(E-2,048)^2/2,048$$

により求めた「適合度値」30,000個の度数分布を示す．「期待度数」の欄には，確率1/2の二項分布から直接計算した，正確な「期待度数」を記した．尚，参考のため，最右欄に自由度1のカイ二乗分布から求めた近似的な期待度数を掲げておいた．

項	χ_1^2 の値	実現度数	期待度数	自由度1の χ^2 分布
1	0 ～ 0.5	15401	15540.5	15615.0
2	0.5～ 1	4680	4711.5	4865.7
3	1 ～ 1.5	3264	3236.2	2899.2
4	1.5～ 2	1871	1860.1	1901.2
5	2 ～ 2.5	1250	1215.8	1303.6
6	2.5～ 3	982	950.6	917.5
7	3 ～ 3.5	601	596.3	656.9
8	3.5～ 4	472	472.8	476.1
9	4 ～ 4.5	407	369.0	348.2
10	4.5～ 5	279	283.6	256.4
11	5 ～ 5.5	204	214.5	189.9
12	5.5～ 6	111	124.3	141.3
13	6 ～ 6.5	105	98.8	105.6
14	6.5～ 7	96	77.8	79.1
15	7 ～ 7.5	56	60.7	59.4
16	7.5～ 8	52	47.0	44.8
17	8 ～ 8.5	47	36.1	33.8
18	8.5～ 9	22	19.1	25.6
19	9 ～ 9.5	34	22.7	19.4
20	9.5～10	23	17.0	14.7
	10 ～10.5	11	8.9	11.1
21	10.5～11	10	10.3	8.5
	11 ～11.5	4	5.3	6.5
	11.5～12	2	4.3	4.9
22	12 ～12.5	1	4.9	3.8
	12.5～13	4	2.5	2.9
	13 ～13.5	1	2.0	2.2
	13.5～14	0	1.6	1.7
	14 ～14.5	1	1.2	1.3
	14.5～15	5	1.0	1.0
	15 ～15.5	2	.8	.8
	15.5～16	0	.6	.6
	16 ～16.5	0	.5	.4
	16.5～17	0	.4	.3
	17 ～17.5	1	.3	.3
	17.5～18	0	.2	.2
	18 ～18.5	1	.2	.2
	18.5～19	0	.1	.1
	19 ～19.5	0	.1	.1
	19.5～20	0	.1	.1
	20 ～	0	.2	.2

確率 1/2 の二項分布に従う」という仮説の下で得られる分布に非常に良く似ていることがわかる．因みに，この度数分布を 22 項から成る多項分布（表 1 の最左欄に従ってクラスをまとめる）と見なして，χ^2 適合度検定を行なうと，その結果

$$\chi^2 = 27.1 \quad (\text{自由度 } 21)$$

を得た．これらのことから，短いスパンでの偶数出現率も良好な性質を持っているように思われる．

3.3.3 偶数・奇数出現の独立性

独立性については，この回路の場合に限れば，相隣り合う乱数間の関連度を測定すれば充分であろう．算術乱数では必ずしも言えないことであるが，物理乱数では回路の構成上何らかの意味で直接の関連が生じそうな組合せのみを考慮すれば良い．例えば「前の前」の乱数との間の関連性は，もし存在したとしても，直接的な「前」の乱数との間の関連性より遙かに小さいと考えられる．

16 進乱数としての独立性は，100 万個程度の比較的少数の乱数を用いた実験で見る限り，充分良好に保たれているように見える．より多数の乱数を用いる実験は，もし独立性に問題があるとしてもその検出には 10 億のオーダの乱数が必要と思われ，時間的制約とパーソナル・コンピュータの信頼性（100 時間を越す連続実験）から実行しなかった．

ここでは，主に回路構成（論理的構成，基板上の配置および配線などの実装上の問題，IC 等の電子部品の選択と性能，電源などからの不要雑音対策など）の不備から生ずる可能性のある独立性の崩れを検出することをねらいとして，偶奇性についてのみ計測を行なっている．この種の崩れは，カウンタの最下位ビットにのみ現われると考えられるからである．

ひとつ前の乱数と現在の乱数の各組合せについて，それらを偶奇性により 4 通りに分類したとき，独立性が保たれていれば，各分類の出現確率は相等しい．その等出現確率性の尺度である χ_3^2 (3.3.1④) の度数分布を表 2 に示した．最右欄の「期待度数」とは，「各分類の出現は

表 2.　偶数・奇数出現の独立性を検証するためのカイ二乗値の分布

発生した乱数列の連続する 4,096 個ごとに，「偶数・偶数」「偶数・奇数」「奇数・偶数」「奇数・奇数」の各組合わせの出現数 (P_0, P_1, P_2, P_3) から

$$\chi_3^2 = \sum_{i=0}^{3} (P_i - 1,024)^2 / 1,024$$

により求めた「適合度値」30,000 個の度数分布を示す．最右欄に自由度 3 のカイ二乗分布から求めた期待度数を掲げておく．

項	χ_3^2 の値	実現度数	期待度数	項	χ_3^2 の値	実現度数	期待度数
1	0～ 1	6044	5962.4	14	13～14	56	51.9
2	1～ 2	6840	6865.4	15	14～15	24	32.7
3	2～ 3	5406	5423.4	16	15～16	20	20.5
4	3～ 4	3917	3904.8		16～17	11	12.8
5	4～ 5	2760	2690.0	17	17～18	3	8.0
6	5～ 6	1714	1805.6		18～19	5	5.0
7	6～ 7	1225	1191.4		19～20	4	3.1
8	7～ 8	738	776.6		20～21	2	1.9
9	8～ 9	509	501.6		21～22	6	1.2
10	9～10	293	321.7		22～23	0	.7
11	10～11	205	205.2		23～24	1	.5
12	11～12	141	130.3		24～25	0	.3
13	12～13	76	82.4		25～	0	.5

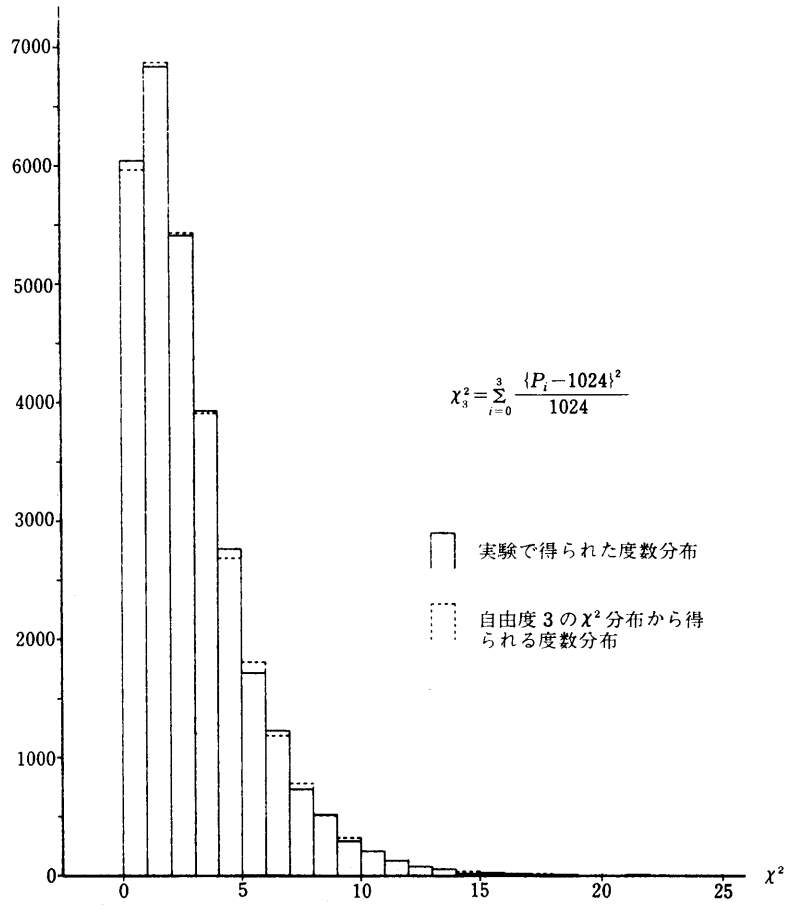

図6 乱数4,096個ごとに求めた χ_3^2 値の度数分布

独立が等確率」との仮説の下で，χ_3^2 が近似的に従う自由度3の χ^2 分布から計算した期待度数値である．仮説の下での正確な χ_3^2 の分布を求めることも可能であるが，標本誤差と比較すれば，この場合の近似による誤差は充分小さい．

表2を見ると，χ_3^2 の分布は自由度3の χ^2 分布と良く似ているように思える．その再確認のために，この度数分布を17項から成る多項分布（表2の最左欄に従う）と見なして χ^2 適合度検定を行なう．計算により，

$$\chi^2 = 17.5 \qquad \text{(自由度 16)}$$

を得，仮説の成立を支持する結論が導かれた．

3.3.4 16進一様性

表3の「物理乱数・μ-com」の欄に示したのは，乱数4,096個ごとに求めた χ_{15}^2 (3.3.1⑤) の度数分布表である．確率変数 χ_{15}^2 は，「乱数が独立で [0, 15] 上の離散一様分布に従う」との仮説の下で，近似的に自由度15の χ^2 分布に従う．この場合にも，近似の精度は充分で，「正しく自由度15の χ^2 分布に従う」として実際上問題ない．

ヒストグラム（図7）を見ると，実験値は仮説に基く期待度数分布に良く従っているように感じられる．念のため，度数を表3の最左欄に従ってまとめ，36項から成る多項分布と考えた

表 3.　代表的な発生法により作られた16進一様乱数列の短い構成単位の統計的性質

表中の数字は，16進一様乱数をそれぞれの方法で発生し，連続する4,096個ごとの各値（0から15まで）の出現回数 $H_i(i=0, 1, \cdots, 15)$ から

$$\chi_{15}^2 = \sum_{i=0}^{15} \{H_i - 256\}^2/256$$

により求めた「適合度値」30,000個の度数分布を示す．最右欄は自由度15のカイ二乗分布から求めた期待度数である．

テストされた一様乱数発生法は，以下に掲げた4類7種である．

乗算合同法：（A）乗数＝39,894,229，法＝2^{32}
（B）乗数＝65,539，法＝2^{32}

m 系 列：原始3項式 $x^{521}+x^{32}+1$ に対応（シフトレジスタ長＝521）
（A）サブレジスタ長＝16（×23）or 17（×9），伏見，手塚 [3]
（B）サブレジスタ長＝16（×31）or 24（×1），伏見 [2]

RND 関数：PC 8001 の BASIC の組込み関数

物理乱数：（大型用）統計数理研究所の物理乱数発生装置，石田ら [5]
（μ-com）μ-computer のための物理乱数発生器，本稿

項	適合度	乗算合同法		m 系列		RND 関数	物理乱数		期待度数
		A	B	A	B		大型用	μ-com	
1	0～ 1	0	0	0	1	0	0	0	.0
	1～ 2	1	0	0	0	0	0	1	.9
	2～ 3	8	12	4	10	45	5	12	11.2
	3～ 4	65	48	68	62	72	52	51	55.8
2	4～ 5	187	186	169	174	75	164	169	168.3
3	5～ 6	363	360	363	380	366	385	377	371.4
4	6～ 7	643	637	689	688	625	656	682	662.9
5	7～ 8	945	1026	1041	1089	735	1089	1029	1016.0
6	8～ 9	1334	1397	1432	1374	1158	1414	1337	1388.0
7	9～10	1765	1752	1754	1849	1421	1745	1702	1733.3
8	10～11	2009	1970	2002	1973	1738	2014	2012	2014.4
9	11～12	2240	2210	2240	2166	2015	2186	2127	2207.0
10	12～13	2256	2403	2328	2259	2316	2334	2280	2301.9
11	13～14	2323	2328	2266	2338	2174	2296	2350	2303.1
12	14～15	2233	2170	2169	2215	2394	2279	2269	2223.4
13	15～16	2034	2132	2055	1982	2137	1990	2065	2081.0
14	16～17	1865	1866	1865	1844	1716	1926	1915	1895.6
15	17～18	1701	1632	1628	1704	1982	1638	1691	1686.0
16	18～19	1545	1448	1512	1394	1655	1478	1480	1468.0
17	19～20	1336	1297	1228	1214	1532	1218	1231	1254.0
18	20～21	1065	1002	1069	1026	886	1098	1073	1053.1
19	21～22	836	885	829	888	950	832	834	870.8
20	22～23	695	687	681	726	715	675	698	709.9
21	23～24	568	597	555	564	837	583	613	571.4
22	24～25	456	459	433	438	643	455	457	454.5
23	25～26	347	353	363	364	351	345	357	357.6
24	26～27	302	299	300	293	335	266	286	278.6
25	27～28	256	230	231	226	306	202	260	215.0
26	28～29	155	180	151	183	129	176	172	164.5
27	29～30	114	112	132	130	158	126	113	124.9
28	30～31	99	93	100	103	119	89	96	94.1
29	31～32	71	63	85	85	101	83	74	70.4

表 3. のつづき

項	適合度	乗算合同法 A	乗算合同法 B	m 系列 A	m 系列 B	RND 関数	物理乱数 大型用	物理乱数 μ-com	期待度数
30	32～33	54	36	53	44	87	62	57	52.3
31	33～34	39	36	57	58	44	36	31	38.7
32	34～35	26	25	36	33	28	32	23	28.4
33	35～36	18	21	33	32	56	23	18	20.7
34	36～37	15	9	19	26	29	12	19	15.1
35	37～38	9	11	19	15	14	6	10	10.9
35	38～39	5	5	10	12	0	9	7	7.8
36	39～40	5	5	6	11	28	6	4	5.6
36	40～41	3	5	7	6	14	3	4	4.0
36	41～42	3	4	2	7	0	7	4	2.9
36	42～43	2	0	6	3	14	2	1	2.0
36	43～44	1	5	2	4	0	2	2	1.4
36	44～45	1	4	1	1	0	0	3	1.0
36	45～46	0	0	0	3	0	0	0	.7
36	46～47	1	0	1	1	0	1	2	.5
36	47～48	0	0	1	0	0	0	2	.3
36	48～49	0	0	1	1	0	0	0	.2
36	49～	1	0	4	1	0	0	0	.5
適合度		39.6	36.7	55.8	86.6	982.5	32.2	30.7	(.0)
確率		.273	.389	.014	.000	.000	.605	.676	(1.000)

適合度： この表を，最左欄の『項』に従って，36のクラスの度数分布表に再編する．そして，各欄の度数分布の期待度数分布に対する『適合度』を計算する．この行は，そのカイ二乗値を示す．

確　率： 自由度35のカイ二乗分布の上側確率をしめす．

場合の χ^2 適合度検定を行なう．計算により，

$$\chi^2 = 30.7 \qquad (\text{自由度 } 35)$$

と求まり，実際に良く適合していることがわかる．

3.3.5 他の乱数での比較実験

この論文で用いられている方法が，どの程度「乱数の悪さ」を検出できるものか確かめるために，他のいくつかの乱数発生法についても同様な実験を行なった．

ここで取上げたのは，良く使われている算術乱数として3発生法5種，および大型計算機用に開発された高速発生器から得られた物理乱数である．

① 乗算合同法による発生法

乗数として 39,894,229 (A) および 65,539 (B) を用いた2種（法はともに 2^{32}）を選び，得られる整数乱数の上位4ビットを16進乱数とした．

② m 系列による発生法

長さ521のシフトレジスタを用いた m 系列乱数2種について行なった．伏見・手塚 [3] のプログラムから得られる [0, 1) 乱数を16倍した整数部を使用したもの(A)，および，シフトレジスタの分割法が簡単な形（16ビット×31，25ビット×1）の場合に対応したプログラム（伏見 [2]）から得られる同様の乱数 (B) を対象とした．

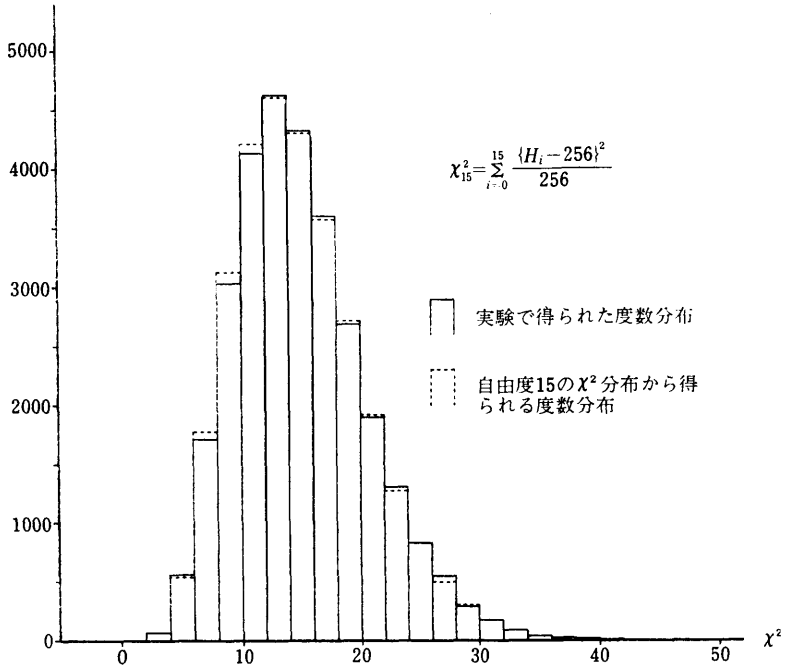

図7　乱数4,096個ごとに求めたχ_{15}^2値の度数分布

③ パーソナル・コンピュータの RND 関数

BASIC に組込まれた乱数発生関数 RND には，乗算合同法を基本とひたものが多いようである．ここで取上げたのは，NEC の PC 8001 に組込まれている RND 関数で，INT (16*RND(1)) により16進乱数を得ることができる．

④ 大型計算機用の高速乱数発生装置

統計数理研究所に設置された HITAC M-200H の周辺機器のひとつである乱数発生装置の出力（20万バイト/秒）の上位4ビットを16進乱数とした．この乱数発生装置については，石田ら [5] および仁木 [9] に記述がある．

表3は，各乱数発生法によって得られた乱数4,096個ごとに，i なる16進値をとった個数を H_i とするとき

$$\chi_{15}^2=\frac{1}{256}\sum_{i=0}^{15}(H_i-256)$$

を計算し，その度数分布を表としたものである．さらに，表3の最下2行には，それらの度数分布が自由度15の χ^2 分布から得られたとしたときの適合度を示す χ^2 値（自由度35）およびその値以上の値をとる確率を示してある．

この表を見る限り，PC 8001 の RND 関数はとても安心して使えるものでないことは一目瞭然であるし，乗算合同法の欠点を克服するものとして考えられた m 系列による発生法も良いとは言えない．一方，乗算合同法の2種は良い成績を残しているが，このテストは乗算合同法の最も得意とする科目について行なったようなものであることを念頭に置くべきである．物理乱数については何も言うことがない．実際，この物理乱数はかなり多種多方面からのテストを

難なくパスしてきた実績を持っている.

4. 結　　語

これまで見てきたごとく，この乱数発生器は，非常に簡単で製作しやすい構造にかかわらず，実際に使用される個数をはるかに越えたサンプル・サイズでのかなり苛酷とも思えるテストに合格するような，良い性質の乱数を発生することがわかった．乱数発生は計算とは並列的に行なうことができるので，パーソナル・コンピュータの計算速度を考えると，発生速度はむしろ速すぎるくらいである.

単にパーソナル・コンピュータ用にとどまらず，この発生器から取り出された乱数を磁気テープ等に記録しておけば，乱数発生装置を持たない大型計算機によるシミュレーションにも充分使用できよう．もちろんこの場合は発生速度を4,000個/秒に落とす必要がある.

発生器の設計を適切に行なえば，手軽に良質の物理乱数が得られる．このような理解が生まれれば，本稿の目的は達せられよう.

なお，査読者の方々から，貴重な助言をいただいた．記して感謝の意を表したい.

参 考 文 献

[1] Arvillias, A.C. and Maritsas, D.G. (1978). Partitioning the period of a class of m-sequences and application to pseudorandom number generation, *J. Ass. Comput. Mach.*, **25**, 675-686.
[2] 伏見正則 (1980). Personal communication (Arvillias and Maritsas (1978) の改良版).
[3] 伏見正則，手塚集 (1981). 多次元分布が一様な擬似乱数列の生成法，応用統計学，**10**, 151-163.
[4] 石田正次 (1965). モンテカルロ法と乱数，科学基礎論研究，**7**, 25-31.
[5] 石田正次，佐藤利男，鈴木亀二郎，下田昭一郎，川瀬哲郎 (1972). ダイオード・ノイズを利用した乱数発生装置，日立評論，**54**, 894-898.
[6] Knuth, D.E. (1981). *The Art of Computer Programming*, Vol. 2 (2d. Ed.), Addison-Wesley, Reading, Mass.
[7] Marsaglia, G. (1968). Random numbers fall mainly on the planes, *Proc. Nat. Acad. Sci. USA*, **61**, 25-28.
[8] Niki, N. (1979). Multi-folding the normal distribution and mutual transformation between uniform and normal random variables, *Ann. Inst. Statist. Math.*, **31**, A, 125-140.
[9] 仁木直人 (1980). 工学的乱数発生，統計数理研究所彙報，**27**, 115-131.
[10] 仁木直人 (1981). 物理乱数，数理科学，**19**, 51-58.

Physical Random Number Generator for Personal Computers

Naoto Niki
(The Institute of Statistical Mathematics)

A physical random number generator has been designed, made and tested by the author. The structure of the generator is very simple so that it may be assembled on a small PC board. The source of a random process is the noise of a Zener diode in a proper bias condition. After being amplified and passed through a high-pass-filter, the noise is applied to a voltage comparator to generate the random pulse process. The number of random pulses (very thin pulses are eliminated) occuring during 125 micro seconds has a distribution closely approximated by the normal distribution with mean between 430 and 470 and variance between 100 and 120. Therefore, the distribution of the least significant digit of the number of pulses in hexadecimal (decimal) representation is very close to the discrete rectangular distribution on [0, 15] ([0, 9]). The maximum relative error of approximation may be estimated at 0.0009 (10^{-8}), theoretically, which may suffice for use on personal computers. The least significant digit of the number of pulses is extracted by use of a hexadecimal (decimal) counter, a Schmitt-trigger gate, a clock circuit and a set of interface circuits between the computer and the generator. The time interval of clock-triggers is selectable between 125 and 250 micro seconds. When the clock is set at 250 micro seconds, the approximation error to the discrete rectangular distribution may become less than 10^{-6} for hexadecimal random numbers or 10^{-16} for decimal ones. Tests for checking the deviation of the empirical distribution from uniformity and the dependence between neighbouring numbers have been made. The χ^2 tests were applied to the even-odd ratio for each group of hexadecimal numbers of size 4,096, the even-odd independence and the uniformity. The empirical distributions of χ^2's, each of 30,000 in size, were compared with the distributions under the hypothesis. The second and third theoretical distributions are well approximated by the χ^2 distributions of degrees of freedom 3 and 15 respectively. The first may be directly calculated. The tests of goodness of fit have shown that the observed numbers have good properties towards true randomness. The numbers derived from m-sequences and the numbers generated by RND function in BASIC have, however, been unfavourable to the hypothesis that they are equidistributed.