数理解析研究所講究録
第371巻 1979年 1-17

二次式と素数

立教大　一般教育部　座間　宣夫

1. 問題の発端.

xを変数とする二次式

$$f_q(x) = x^2 - x + q \qquad (1)$$

のxに自然数を代入してゆくと，適当なqを取ると，$1 \leqq x \leqq q-1$を満足するすべての自然数に対し$f_q(x)$の値が素数となることは古くから知られていた．たとえば，$q = 3$，5，11，41についてこのような現象が起る．ところで，

$$f_{41}(x) = x^2 - x + 41$$

のxに自然数を次々と代入してゆくと，$f_{41}(x)$の値には素数が異常に多く含まれていることが判る．

筆者は（1）の形の二次式のxに自然数を代入してゆくとき，どのようなqに対して，素数を得る可能性が大きくなるかについて実験を試みた．

2

まずxについての二次式のxに自然数を順次代入してゆくとき何が起るかについて考えてみる．実例として$f_1(x)$を$x=1$から$x=11$まで計算して素因数分解し次表にかかげる．

x	$f_1(x)$
1	1
2	3
3	7
4	13
5	$21=3\times7$
6	31
7	43
8	$57=3\times19$
9	73
10	$91=7\times13$
11	$111=3\times37$

$f_1(2)$の値3とxの値2とを加えて得られる5をxに代入すると，$f_1(5)=21=3\times7$がえられる．さらに$f_1(5+3)$，$f_1(8+3)$は3の倍数であり，$f_1(5+7)$は7の倍数である．この事実を定理の形にまとめると次のようになる．

[定理]　qを自然数とすれば，

$$q \mid f_p(x) \longleftrightarrow q \mid f_p(x+q)$$

[系]　$f_p(x)$が素因数rをもてば，rは$f_p(x-r)$かまたはその素因数になる．

上記の定理で述べた性質をquasi-Erathostenes性（略してQER性）と呼ぶことにする．自然数列から素数を検出するにはエラトステネスのふるいが使われるが，$\{f_q(x) \mid x=1, 2, \cdots\}$から素数を検出するにはQER性が使用される．ただし，後者の場合にはしるしがつけられなくてもいちおう素因子分解をしないと素数であるかどうか判らないのである．

2. 実験の内容．

QER性によれば，素数の生ずる可能性の大きい$f_p(x)$は，xが小さい自然数を取るとき素数となる割合が大きいもの，また素因子がなるべく大きくなるようなものであろうという推定が考えられる．この推定にもとずき，実験は次の二段階にわけて行なうことにした．

（１）p を 128189 以下の素数とし，$f_p(1)$ から $f_p(100)$ を計算し，その中に含まれる素数の個数ならびに素因数をしらべる．

（２）段階（１）の結果，素数の個数が７０個以上のものならびに素因子の大きいものについて，さらに大きい x の値に対して $f_p(x)$ を計算してその中に含まれる素数の個数をしらべる．

実験（１）の結果をまとめると，次の表が得られる．

f_p中の素数の数 ＼ p	1番目～1300番目の素数（大体1万まで）	1301番目～2300番目の素数（大体2万まで）	2301番目～4200番目の素数（大体3万まで）	3201番目～4200番目の素数（大体4万まで）	4201番目～5100番目の素数（大体5万まで）	5101番目～6000番目の素数（大体6万まで）	6001番目～6900番目の素数（大体7万まで）
0～4	1						
5～9	4	5	7	11	14	8	11
10～14	43	69	66	83	84	90	84
15～19	137	127	137	156	163	155	183
20～24	188	173	171	196	185	185	190
25～29	221	198	168	203	156	173	157
30～34	208	160	133	138	126	121	117
35～39	178	111	89	101	72	81	75
40～44	107	70	65	49	58	51	42
45～49	80	42	29	33	25	28	21

表1

50~54	44	22	15	13	10	3	12
55~59	32	8	8	12	3	7	5
60~64	22	7	7	5	1	4	2
65~69	11	5	3	1	1	1	1
70~74	4	3				1	
75~80	1		1				

f_p中の素数の数 ＼ p	6901番目~7800番目の素数（大体8万まで）	7801番目~8700番目の素数（大体9万まで）	8701番目~9500番目の素数（大体10万まで）	9501番目~10400番目の素数（大体11万まで）	10401番目~11300番目の素数（大体12万まで）	11301番目~12000番目の素数（128189まで）
0~4						
5~9	14	7	9	20	15	16
10~14	90	93	77	95	105	86
15~19	163	166	168	198	195	134
20~24	192	202	178	197	208	147
25~29	160	161	131	153	147	120
30~34	121	119	106	108	94	89
35~39	79	77	60	68	65	51
40~44	48	34	40	37	39	33
45~49	19	26	20	13	20	15
50~54	8	12	6	6	7	6
55~59		3	2	5	4	2
60~64		1	2		1	1
65~69	1		1			

表1.

6

上の表は，$f_p(1)$ から $f_p(x)$ の中に含まれる素数の個数を $A_p(x)$ と書くことにすると，$A_p(100)$ の値の分布を示す．$A_p(100)$ が70以上になる p の値を次表に示す．

p	$A_p(100)$
41	86
1997	70
3461	70
3917	70
4931	73
7517	73
8081	72
9281	71
11777	70
14627	71
19421	74
21377	75
27941	77
55661	73

また $f_p(x)$ の中の100以下の素因数が大きく，かつ個

数の少ないものを次表にかかげる.

〔表2〕

p	100以下の素因数
55661	43, 53, 67.
27941	47, 67, 73, 79, 83, 89.
21377	37, 61, 71, 73, 79, 89, 97.
*114467	41, 43, 47, 53, 61, 71, 83, 97.
*41	41, 43, 47, 53, 61, 71, 83, 97.
19421	47, 53, 59, 67.
7517	29, 53, 67, 71, 73, 79, 97.
14627	37, 41, 43, 47, 59, 79, 89, 97
*11777	17^2, 41, 43, 47, 53, 61, 71, 83, 97.
8081	31, 43, 53, 59, 73, 79, 89, 97.
3461	23, 31, 53, 59, 61, 67,

	71, 73
*4931	11^2, 41, 43, 47, 53, 61, 71, 83, 97.
9281	37, 41, 47, 67, 71, 83, 89.
3917	29, 31, 37, 41, 59, 71.
*1997	7^2, 41, 43, 47, 53, 61, 71, 83, 97.

上の表を見て気付くことは，先頭に＊をつけた行の素因数は先頭の数を除けば，全く一致している．41と114467の場合は完全に一致している．試みにこれらの数の判別式を計算してみると，

$$1-\quad 41\times 4=-163$$
$$1-114467\times 4=-163\times 53^2$$
$$1-\quad 11777\times 4=-163\times 17^2$$
$$1-\quad 4931\times 4=-163\times 11^2$$
$$1-\quad 1997\times 4=-163\times 7^2$$

となる．この事実が$f_p(x)$中の素数の位数にどのような影響を与えるかについては4節でのべることにする．

3. $A_p(x)$ の値

前節の内容はいわば予備実験に相当するが，その結果にもとずいて $f_p(1)$ より $f_p(x)$ の中に含まれる素数の個数の計算を行なった．$f_p(x)$ の計算をしてからその値の素因数分解を行なっていたのでは計算に時間がかかって実用にならない．そこで，QERを利用し，$f_p(a)$ の素因数を r とすると，$a+r$，$a+2r$，$a+3r$，…にしるしをつけ，しるしのついた x の値に対して $f_p(x)$ は素数とならないことを利用して計算時間の節約をはかった．この計算の結果得られた $A_p(x)$ の値は次のとうりである．

p	$A_p(100)$	$A_p(10000)$	$A_p(20000)$	$A_p(30000)$
41	86	4149	7620	10842
1997	70	3561		
3461	70	3760		
3917	70	3646		
4931	73	3715		
7517	73	3944		
8081	72	3853		
9281	71	3650		
11777	70	3887		

14627	71	3893		
19421	74	4024		
21377	75	4255		
27941	77	4466	8258	
55661	73	4544	8458	12191
114467	62	4160	7695	

詳細な数値の表は筆者による文献（２）を参照されたい．

この結果を要約すると次のようになる．$x=1$ から $x=100$ までで考えれば，$f_{41}(x)$ 中に含まれる素数の個数が圧倒的に多いが，x の値をふやしてゆくと，f_{41} よりもっと多くの素数が含まれるものが出てきて，特に f_{55661} に著しく多くの素数が含まれることが判る．

$A_{41}(100)$ が大きいという事実は別の解釈を与えることが出来る．$A_p(x)$ と $f_p(x)$ の判別式

$$\Delta = 1-4p$$

の間には密接な関係がある．すなわち，Q を有理数体とするとき，$Q(\sqrt{\Delta})$ のイデアル類数が１となることが，$f_p(1)$ から $f_p(p-1)$ までのすべての値が素数となるための必要十分条件である．ところが，類数が１となる虚二次体は Δ が

$$-1,\ -2,\ -3,\ -7,\ -11,\ -19,\ -43,$$
$$-67,\ -163$$

となる場合だけである．つまり，$f_p(1)$ から $f_p(p-1)$ のすべてが素数となるのは，p が

$$2, 3, 5, 11, 17, 41$$

となる場合に限られることになる．すなわち，前の実験の結果はこのような場合以外に，素数の出現回数の多い $f_p(x)$ が存在することを意味する．

4. Hardy, Littlewood の予想.

本節の内容は筑波大学内山三郎教授が筆者の実験結果について与えられた注意である．

Hardy と Littlewood は文献［1］において，二次式 an^2+bn+c に1からNまでの自然数値を与えたとき，その中に含まれる素数の個数の近似公式を与えた．筆者の実験は，二次式 x^2-x+p について行なわれているが，この場合にこの近似公式をあてはめてみると次のようになる．

$$A_p(N) \sim c_p \int_2^N \frac{du}{\log u} . \qquad (1)$$

ただし，c_p は次の式で定義される定数である．

$$c_p = \prod_{q:\,3\text{以上の素数}} \left(1-\left(\frac{1-4p}{q}\right)\frac{1}{q-1}\right) \qquad (2)$$

（2）において，$\left(\frac{1-4p}{q}\right)$ はLegendre の記号である．（1）から判るように，$A_p(N)$ すなわち二次式中に含まれる素数の数は c_p の大きさに依存するが，$p=41$ とおくと

$$1-4\times 41=-163$$

となり，$p=114467$ とおくと，

$$1-4\times 114467=-163\times 53^2$$

となって，

$$c_{114467}=\frac{52}{51}c_{41} \qquad (3)$$

となる．筆者の実験の結果によると，$p=41$ の場合と $p=114467$ の場合をくらべると，後者の方に素数の出現回数が多いことと（3）とが完全に符合する．

以上が内山教授の御指適である．実験の結果によると，

$$\frac{A_{114467}(20000)}{A_{41}(20000)}=\frac{7695}{7620}\fallingdotseq 1.010$$

で $\frac{52}{51}\fallingdotseq 1.020$ よりやや少ないが，この値をもっとふやしてみないと確実な結果は判らない．

実は，これと似た現象は筆者の他の実験でも得られており，2節の末尾で述べたように，$p=11777, 4931, 1997$ の場合に，判別式がそれぞれ -163×17^2，-163×11^2，-163×7^2 となっていた．この場合に上の計算と同様の試算を行なうと次のようになる．まず，

$$c_{11777} = \frac{16}{17} c_{41} \fallingdotseq 0.9411 c_{41},$$
$$c_{4931} = \frac{10}{11} c_{41} \fallingdotseq 0.909 c_{41},$$
$$c_{1997} = \frac{6}{7} c_{41} \fallingdotseq 0.857 c_{41}.$$

となるが，筆者の実験によれば，

$$\frac{A_{11777}(10000)}{A_{41}(10000)} = \frac{3887}{4149} \fallingdotseq 0.937,$$

$$\frac{A_{4931}(10000)}{A_{41}(10000)} = \frac{3715}{4149} \fallingdotseq 0.895,$$

$$\frac{A_{1997}(10000)}{A_{41}(10000)} = \frac{3561}{4149} \fallingdotseq 0.858$$

となって，相当程度一致すると考えられる．このような方法で，Hardy, Littlewood の近似公式の正当性についての資料がえられたことはきわめて興味深い．

5. 実験によって生じた問題点

筆者の実験の結果からいくつかの疑問が生ずる．それらを以下列挙しよう．

（1） 2節の表1によると $A_p(100)$ は p の増加に伴なってだんだん減少の傾向を示す．この事実は p がもっと大きくなったときも依然成立つであろうか？

（２）２節の表２は最初の１２０００個の素数pに対して，$f_p(x)$の１００以下の素因数をしらべた結果得られたもので，１００以下の素因数の個数がこの表にのせられたものより少ないものはないし，また最小の素因数がこの表にのせられたものを越えるf_pはいままでのところみつかっていない．pをどんどん増加させていくと，f_pの最小の素因数がいくらでも大きくなるように出来るであろうか？ 例えば，１００以下の素因数が一つもないようなf_pは果して存在するのであろうか？

６．電子計算機に対する整数論の実験についての一般的注意．

最後に，整数論の実験の経験からみて，整数論の実験を行なうのに適した電子計算機とそのオペレーティング・システムについて若干の感想を述べてみたい．

そのような電子計算機が整数の計算に必要な演算機能を持ち，そのデータ構造は整数の計算に適していなくてはならないのは当然なことである．これら二つのうち，データの構造の方が演算機能よりもっと多くの問題を含んでいる．例えば現在までもっとも世界で広く使われた電子計算機であるＩＢ

M社のS／３６０では，整数は１６ bitの２進数として表わされ，十進数で-３２７６８から３２７６７までしか表わせない．これでは整数論の実験に必要な大きさを持っているとは言えず，少なくとも整数は３２ bit以上の２進数でないと十分でない．筆者が実験に使用した計算機も通常の整数は２進数１６ bitで，２進数３２ bitの倍精度整数を使うことは出来るが，倍精度整数の演算を行なう演算装置を持たず，単精度整数二つを使って倍精度整数と考え，プログラムで演算を行なっている．この方法によると，計算時間が長くなってしまい，単位時間当りの計算コストの増大をまねくことになる．大規模な計算を行なうためにもっとも重要な要素は，単位時間当りの計算コストと使い易さにあるといっても過言ではないと考える．

実数値を使う数値計算を行なう電子計算機では，数はすべて浮動小数点を使って表わされていて，ちょっと考えてみると整数論の計算と両立しないようにみえる．しかし，例えばCDCのCyberシリーズ計算機のように，２進数６４ bitを基本にしている計算機では，仮数部４８ bitを使えば整数の計算に十分使うことが出来る．新しい計算機を作るのが大変ならば，２進数６４ bitの演算を行なう演算装置を別に作っておき，従来の計算機をデータのやりとりを制御するフロン

ト・エンド・プロセサーとして使って，その制御のもとに使えば十分である．要するに，現在市販されている計算機の中には整数の計算に適していないものが多いが，多少の改良で十分整数論の計算に使えるように出来るうえに，実数による数値計算に向く計算機としても両立しうると考える．

さて，整数論実験の計算には大別して次の二つがある．

（1）メルセンヌ数の発見のように，一つのプログラムで大規模な計算を行なうもの．

（2）筆者の実験のように小規模な計算を多数回繰返し，その結果を編集して最終的な結果とするもの．

例えば，2節の表1は，各pに対して，$A_p(100)$や100以下の素因数を記録した磁気テープからプログラムを使って計算機で編集して作られた．（1）の型のプログラムは，そのプログラムを実行する能力を持つ電子計算機があればよいが，（2）の型のプログラムでは，磁気テープや磁気ディスクに記録されたデータの集積，いわゆるファイルの処理がうまく出来ないと困る．筆者の推定では四色問題の電子計算機による解決のさいも，（2）の型の業務によったものと考える．この型の業務ではファイルから所要の結果をうるためのプログラムが容易に作られないと，作業の能率は上らない．要約すると，整数論の計算でも，

（1）ファイル中心，

（2）データ・ベース指向

のオペレーティング・システムが必要になったといえる．このような業務では，事務計算と同様にシステム設計の手法によるシステム作製が必要なことはいうまでもない．この問題についてはまた稿をあらためて論じてみたい．

［参考文献］

（1）G. H. Hardy, J. E. Littlewood.
'Some problems of 'Partitio Numerorum'; Ⅲ: On the expression of a number as a sum of primes.'
Acta Mathematica 44 (1922), pp 1—70.

（2）座間宣夫.
'二次式と素数 —電子計算機による計算調査の結果—'.
St. Paul's Review of Science, Vol. 4, No. 1, pp 11—36, 1979.