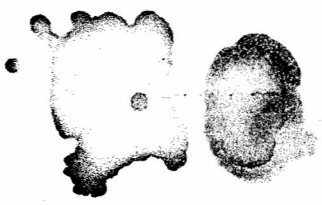
MATSCIENCE REPORT 87.

Proceedings of the Conference on

MATRIX ALGEBRA, COMPUTATIONAL METHODS AND NUMBER THEORY



MYSORE

September 6-9, 76.

THE INSTITUTE OF MATHEMATICAL SCIENCES. MADRAS-600 020. (India)

January, 1977.

Proceedings of the

Conference

on

MATRIX ALGEBRA, COMPUTATIONAL METHODS AND NUMBER THEORY

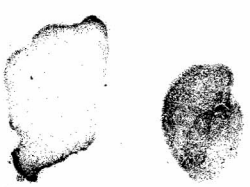
held at

The Institution of Engineers (India), Mysore

(6th to 9th September, 1976)

Convener

Professor Alladi Ramakrishnan
Director, Matscience, Madras.



Edited by

Professor N.R. Ranganathan
MATSCIENCE, Madras.

THE INSTITUTE OF MATHEMATICAL SCIENCES
Madras-600020. India.

FOREWORD

Conference on Matrix Algebra, Computational Methods and Number Theory was organised by Matscience and was held for four days from 6th to 9th September, 1976 at the Institution of Engineers, Mysore. About thirty participants including post-graduate teachers and under-graduate students of promise took part in the conference. The academic members of the Institute as well as a few scientists from different institutions in India lectured on various topics relating to the theme of the conference.

Inaugurating the conference Professor Alladi Ramakrishnan, Director of Matscience, observed that these conferences conducted by the Institute every year afford a forum for presentation of original ideas by the younger generation of scientists. In this conference, work at Matscience relating to new connection between number theory and matrix algebra established by two young scientists Dr.R.Jagannathan (now at St.Joseph's College, Trichy) and Mr.Krishnaswami Alladi (University of California at Los Angeles) was presented.

The organisers of this conference wishes to thank all the participants for their cooperation which contributed to the success of the conference.

The Editor wishes to place on record the help rendered by Mr.R.Jayaraman, Mr.R.Ganapathy and Mr.D.Varadarajan in the organisation of the conference. He also wishes to mention with thanks the effort of Mr.R.Jayaraman in bringing out the proceedings of this conference as a Matscience Report.

N.R.Ranganathan
Editor

CONTENTS

A NEW LOOK AT PERMUTATIONS, PARTITIONS AND L-MATRICES

Alladi Ramakrishnan,
MATSCIENCE, The Institute of Mathematical Sciences,
Madras, INDIA.

+*+*+*+

The old order changeth, yielding place to new
And God fulfils himself in many ways
Lest one good custom should corrupt the world

-Tennyson

The concept of permutation is so familar that one is convinced it is futile to explore this 'stale, unprofitable' domain. However we now show that fascinating features are revealed when we take a new look simultaneously at permutations, partitions and L-matrices.

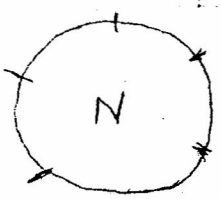
A permutation is an arrangement or a rearrangement of n distinguishable objects according as we take a 'passive' or an 'active' interpretation of the concept. The two are equivalent but it is convenient to adopt the 'active' approach since it leads naturally to the representation of permutations by matrices and arrangements as vectors.

If there is an arrangement of N distinguishable objects, we call a permutation an N-cycle or a cycle of length N if it shifts the positions of <u>all</u> the N objects. A typical cycle is denoted this by

$$1 \longrightarrow S_1 \longrightarrow S_2 \cdots\cdots \longrightarrow S_{N-1} \longrightarrow S_{N-1} \longrightarrow 1$$

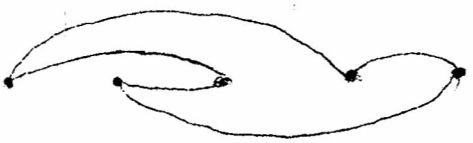
the first object goes to S_1 theplace, the S_1th to the S_2th,.... and the S_Nth to the first place where the S_i take the values 2,3,....N-1, no two assuming the same value.

Thus there are (N-1)! distinguishable cycles of length N. It is appropriate to represent by a cycle by a circle with its length denoted inside

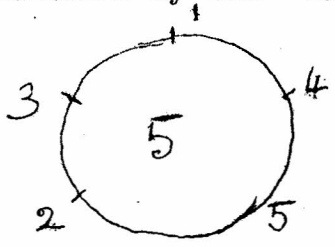


For example taking 5 objects a typical cycle which can be graphically represented as

$1 \longrightarrow 4 \longrightarrow 5 \longrightarrow 2 \longrightarrow 3 \longrightarrow 1$



can still be represented by the 'topologically equivalent' circle



It is important to note that in our new approach, (N-1)! cycles represent the permutation of N objects in which <u>all the objects are shifted and each only once</u>, and does not represent the permutations of (N-1) objects as is customarily understood.

A permutation in general can be broken up into cycles. If

$$N = k_1 n_1 + k_2 n_2 \ldots\ldots k_r n_r$$

a typical permutation will have k_1 cycles of length n_1, k_2 cycles of length n_2 ..... and k_r cycles of length n_r corresponding to the <u>partition</u> of the number N into integers, n_i occurring k_i times ($i = 1,\ldots.r$) and $r = (1,\ldots.N)$.

We shall call the integers n_i <u>the components</u> of the partition of N. The number of distinguishable permutations corresponding to this partition is

$$R(N,n_i,k_i,r) = \frac{N!}{(n_1!)^{k_1}(n_2!)^{k_2}\ldots.(n_r!)^{k_r}} \cdot \frac{(n_1-1!)^{k_1}(n_2-1!)^{k_2}\ldots.(n_r-1!)^{k_r}}{k_1!\,k_2!\ldots.k_r!}$$

$$= \frac{N!}{n_1^{k_1} n_2^{k_2} \ldots. n_r^{k_r}.k_1!\,k_2!\ldots k_r!}$$

This is obtained by counting the number of ways in which the partition of the N objects can be done, computing the number of cycles and correcting for the number or cycles equal length.

The summation of the above expression over all k_i and n_i and r yields the total number of permutations, N!

$$\sum_{k_i,n_i,r} R(N,r,|\ n_i,k_i) = 1$$

The denominator of R is this <u>the product of the components of the partition and the factorials of the exponents</u>.*

An N-cycle can be represented by the matrix

$$C = \begin{bmatrix} & 1 & & \\ & & 1 & \\ & & & \ddots \\ 1 & & & \end{bmatrix}$$

The (N-1)! cycles of length N can be obtained by changing the positions of all the columns taking care that the diagonal is always null. The eigenvectors of the C-matrix are

*The connection between the sum of the components and the product of the components has been studied in detail by K.Alladi and also K.Alladi and P.Erdős when the components are primes.

$$\begin{matrix} 1 & 1 & & 1 \\ 1 & \omega & & \omega^{N-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & \omega^{N-1} & & \omega^{(N-1)} \end{matrix}$$

corresponding to the eigenvalues $1, \omega, \omega^2 \ldots, \omega^{N-1}$ where $\omega = e^{2\pi i/N}$ is the primitive root of unity. The effect of any one of the $(N-1)!$ cyclic matrices is to shift the positions of all the elements in the vectors.

In L-matrix theory we define the matrix

$$B = \begin{bmatrix} 1 & & & & \\ & \omega & & & \\ & & \omega^2 & & \\ & & & \ddots & \\ & & & & \omega^{N-1} \end{bmatrix}$$

and recognise the fundamental commutation relation

$$CB = \omega BC$$

which is the basis of the entire theory. If we move the columns of C according to an N-cycle the corresponding changes can be made in B to preserve the ω-commutation relation. It is only for notational and conceptual convenience we arrange the roots in ascending powers in the eigen vectors and define the N-cycle corresponding to C as

$$1 \longrightarrow N \longrightarrow N-1 \quad \ldots \quad 2 \longrightarrow 1$$

Any one of the $(N-1)!$ cycles of length N is as good as another provided we make the same changes in B to preserve the ω-commutation relation.

In the traditional concept of treating a permutation as a product of interchanges an object may be subject to many interchanges while according to our approach in any permutation an object is shifted only once in any cycle of length $n > 1$ or not at all in the degenerate cycle of length 1.

We illustrate the theorem enunciated in the particular case of 7. Corresponding to the fifteen partitions 7, (6,1), (5,2) (5,1,1), (4,3), (4,2,1,), (4,1,1,1), (3,3,1), (3,2,2), (3,2,1,1), (3,1,1,1,1), (2,2,2,1) (2,2,1,1,1,1), (2,1,1,1,1,1,1), (1,1,1,1,1,1,1) the number of permutations are 720, 840, 504, 504, 420, 630, 210, 280, 210, 420, 70, 105, 105, 21, 1 which add upto $7! = 5040$.

REFERENCES

Alladi Ramakrishnan, 'L-Matrix theory or The Grammar of Dirac Matrices' Tata McGraw Hill Pub. Co.,1972.

Alladi Ramakrishnan, 'A Matrix Decomposition Theorem' Journal of Mathematical Analysis and Applications, Vol.40, No.1.,pp.36-38,1972.

A MATRIX APPROACH TO CERTAIN NUMBER THEORETIC PROBLEMS

R.Jagannathan
MATSCIENCE, The Institute of Mathematical Sciences
Madras-600020. INDIA.

****

Abstract: Certain fundamental results of number theory are understood in terms of a matrix approach.

...

The main purpose of this paper is to describe a simple matrix theoretic formulation of certain basic number theoretic problems. My personal realization of this approach has its origins in a paper by me and Ranganathan dealing with some group theoretic aspects of generalized Clifford algebras studied in the context of L-matrix theory developed by Alladi Ramakrishnan and his collaborators [1]. During the course of the work reported in [2] we had to conjecture a number theoretic identity, which was later shown by Krishnaswami Alladi [3] to be a simple consequence of the famous Mobins inversion formula. Only then I realized that the Mobius inversion can be simply understood as a matrix inversion. Here I will be describing the consequences of such a simple approach (For a sophisticated version of this approach see (4)

Let me recall first a fundamental theorem of matrix theory. If

i) $AB = C$ (1)

where (A,B,C) are all $m \times m$ matrices a unique solution for B exists as

$$B = A^{-1}C, \quad AA^{-1} = A^{-1}A = I. \tag{2}$$

provided

$$\det A \neq 0 \tag{3}$$

Now let me consider matrices of a special structure. An arithmetic function $f(n)$ is defined only for integers $n \geq 1$. Associated with any such arithmetic function $f(n)$ let me construct a lower triangular matrix F with elements.

$$F_{nd} = \begin{cases} f\left(\frac{n}{d}\right) & \text{if } d \text{ divides } n \ (\text{if } d/n) \\ 0 & \text{if } d \text{ does not divide } n \ (\text{if } d \nmid n) \end{cases} \tag{4}$$

It is easy to verify that products of such matrices are also of the same type and the product is commutative. Then let (F,G,H) be three such matrices associated respectively with three arithmeti functions (f,g,h). If F and H are known and G is unknown and

$$FG = H \tag{5}$$

then a unique solution for G exists as

$$G = F^{-1}H \tag{6}$$

provided

$$\det F \neq 0 \quad \text{(e.g. } L(1) \neq 0) \tag{7}$$

It will be seen that when (7) is satisfied F^{-1} is also of the same structure (4). (i.e. $F^{-1}_{nd} = 0$ if $d \times n$, $F^{-1}_{nd} = F^{-1}_{n_1 d_1}$ $= f^{-1}\left(\frac{n}{d}\right)$, $\forall\ k \geq 1$). Hence the solution (6) provides the unknown g in terms of f and h. To this end let me just calculate F^{-1} explicitly assuming that F^{-1} is of the structure (4). Thus

$$g) \quad F^{-1}F = I \tag{8}$$

implies

$$g) \quad \sum_{d|n} f^{-1}\left(\frac{n}{d}\right) f(d) = \delta_{n1} \tag{9}$$

Then, it follows by simple calculations that

$$f^{-1}(1) = \frac{1}{f(1)} \tag{10}$$

and $\forall\ n > 1$

$$f^{-1}(n) = \sum_{k=1}^{\Omega(n)} \frac{(-1)^k}{f(1)^{k+1}} \left\{ \sum \left(f(d_1) f(d_2) \cdots f(d_k) \,\middle|\, \begin{matrix} d_1 \cdots d_k = n \\ d_1 \cdots d_k > 1 \end{matrix} \right) \right. \tag{11}$$

where $\Omega(n)$ is the Ramanujan function defined by

$$\Omega(n) = \sum_{i=1}^{r} \alpha_i \quad \text{when } n = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \tag{12}$$

Distinct primes are denoted by p_i . In (11) the summation includes all possible permutations of d_i's. also Equations (10-11) represent a simple and explicit version of Ward's theorem (cf(4)) (Reonick [5] has given a sophisticated version of the same) Introducing ordering among the d_i's (ii) can be written as

$$f^{-1}(n) = \sum_{k=1}^{\Omega(n)} \frac{(-1)^{k}}{f(1)^{k+1}} \left\{ \sum \frac{k!}{k_1! \cdots k_s!} \prod_{i=1}^{s} f(d_i)^{k_i} \;\middle|\; \begin{array}{l} \prod_{i=1}^{s} d_i^{k_i} = n \\ \sum_{i=1}^{s} k_i = k \\ 1 < d_1 < \cdots < d_s \end{array} \right\} \quad (13)$$

Thus I have arrived at the values of the elements of F^{-1} explicitly assuming that it has the structure (4). The assumption is justified and the correctness of (13) is affirmed by the uniqueness of the inverse of a nonsingular matrix such as F. Hence the equations (4,5,6,7,13) lead to the result that

$$\text{if } \sum_{d/n} f\left(\frac{n}{d}\right) g\left(\frac{d}{\ell}\right) = h\left(\frac{n}{\ell}\right); \; f(1) \neq 0 \quad (14)$$

then

$$g\left(\frac{n}{\ell}\right) = \sum_{d/n} f^{-1}\left(\frac{n}{d}\right) h\left(\frac{d}{\ell}\right) \quad (15)$$

where $f^{-1}\left(\frac{n}{d}\right)$ is given by (13). The familiar inversion formula is obtained by putting $\ell = 1$ without loss of generality. Mobius inversion formua corresponds to the special case when

$$f(n) = \gamma(n) = 1 \;, \; \forall\, n \quad (16$$

Denoting $\mathcal{L}^{-1}(n) = M(n)$, by (13)

$$M(n) = \sum_{k=1}^{r(n)} (-1)^k \varphi(n,k) \tag{17}$$

where $\varphi(n,k)$ is the number of ways, including permutations, of writing n as a product of k factors. If $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, $\varphi(n,k)$ evidently denotes the number of ways of writing $(\alpha_1, \alpha_2, \cdots, \alpha_r)$ as

$$(\alpha_1, \alpha_2, \cdots, \alpha_r) = (1,1 \quad 1) \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2r} \\ \vdots & & & \\ \alpha_{k1} & \alpha_{k2} & \cdots & \alpha_{kr} \end{pmatrix} \tag{18}$$

with no row of A vanishing. Let $\binom{\alpha_i + k - 1}{\alpha_i} = B(\alpha_i, k)$ denote the number of ways in which one can distribute α_i indistinguishable objects in k cells (Bore-Einstein statistics). Then the total number of solution for A including the cases of vanishing rows is

$$\varphi(n,k) = \prod_{i=1}^{r} B(\alpha_i, k) \tag{19}$$

Since $\binom{k}{k-t} \varphi(n,t)$ gives the number of A's satisfying (18) with (k-t) vanishing rows

$$\sum_{t=0}^{k} \binom{k}{k-t} \varphi(n,t) \equiv \sum_{t=0}^{k} \binom{k}{t} \varphi(n,t) = \varphi(n,k) \tag{20}$$

$$\varphi(n,0) = \varphi(n,0) = 0 \tag{21}$$

Hence

$$\varphi(n,k_2) = \sum_{t=0}^{k_2} (-1)^{k_2-t} \binom{k_2}{t} \varphi(n,t) \qquad (22)$$

With this substituted in (17) it reads

$$M(n) = \sum_{k_2=1}^{\Omega(n)} (-1)^{k_2} \left[\sum_{t=1}^{k_2} (-1)^{k_2-t} \binom{k_2}{t} \left\{ \prod_{i=1}^{r} \binom{\alpha_i+t-1}{\alpha_i} \right\} \right] \qquad (23)$$

$$\forall n = (p_1^{\alpha_1} \cdots p_r^{\alpha_r}) > 1$$

This complicated-looking function $M(n)$, really has a very simple structure, discovered originally in 1832 by Mobius as

$$M(n) = \begin{cases} 1 & \text{for } n=1 \\ (-1)^r & \text{if } n = p_1 p_2 \cdots p_r \\ 0 & \text{otherwise.} \end{cases} \qquad (24)$$

Then from (14) and (15) the famous inversion formula results as

$$\text{if } \sum_{d|n} g(d) = h(n) \text{ then } g(n) = \sum_{d|n} M\left(\frac{n}{d}\right) h(d) \qquad (25)$$

so far I have described how the Mobius inversion formula (25) can be understood in terms of matrix theory. Now I shall proceed to describe how to understand another well-known basic Mobius inversion formula forgotten since the time of Mobius and Bachmann (cf. 7)) perhaps due to its failure to be applicable. The other well-known formula is that

$$\text{if } \sum_{k_2 \le x} f\left(\frac{x}{k_2}\right) = g(x) \text{ then } f(x) = \sum_{k_2 \le x} g\left(\frac{x}{k_2}\right) M(k_2), \qquad (26)$$

$$x \ge 1$$

To understand this let me take the following **matrix product**

$$(\mathcal{H}N)_{x\ell} = \sum_{k} \mathcal{H}_{xk} N_{k\ell} \tag{27}$$

where $\mathcal{H}$ has a continuous row index x from 0 to ∞ and a discrete column index k from 1 to ∞. The index is also discrete from 1 to ∞. The matrices $\mathcal{H}$ and N are defined as

$$\mathcal{H}_{x\ell} = \begin{cases} f\left(\frac{x}{\ell}\right) & \text{if } x \geq \ell \\ 0 & \text{if } x < \ell \end{cases} \tag{28}$$

$$N_{k\ell} = \begin{cases} \nu\left(\frac{k}{\ell}\right) = 1 & \text{if } \ell / k \\ 0 & \text{if } \ell \nmid k \end{cases} \tag{29}$$

Let M be the matrix inverse to N such that

$$M_{k\ell} = \begin{cases} \mu\left(\frac{k}{\ell}\right) & \text{if } \ell / k \\ 0 & \text{if } \ell \nmid k \end{cases} \tag{30}$$

and

$$MN = NM = I \tag{31}$$

Then (26) is just

$$\text{if } \mathcal{H}N = \mathcal{G} \quad \text{then } \mathcal{H} = \mathcal{G}M \tag{32}$$

where $\mathcal{G}$ is the matrix associated with $g(x)$ and defined analogous to (28). The truth of (32) and hence of (26) follows immediately by noting that

$$\mathcal{H}N = \mathcal{G} \rightarrow \mathcal{H} = \mathcal{H}(NM) = (\mathcal{H}N)M = \mathcal{G}M \tag{33}$$

In fact one can prescribe the matrices $(\mathcal{H}')_{x,\ell}$ and $(\mathcal{G}')_{x,\ell}$ associated with the two functions $f(x)$ and $g(x)$ analoguous to (28) using any binary relationship between x and ℓ as

$$\mathcal{H}'_{x,\ell} = f(x \circledast \ell), \quad \mathcal{G}'_{x,\ell} = g(x \circledast \ell) \tag{34}$$

such that $((x \circledast \ell) \circledast b_2) = (x \circledast \ell b_2)$ and have the inversion formula (32) or

$$\text{if } \sum_x f(x \circledast \ell b_2) = g(x \circledast b_2) \text{ then } f(x \circledast \ell) = \sum_{b_2} g(x \circledast \ell b_2)\mu(b_2) \tag{35}$$

For example one has the inversion formula

$$\text{if } \sum_{b_2} f(x, b_2) = g(x), \text{ then } g(x) = \sum_{b_2} g(x, b_2)\mu(b_2) \tag{36}$$

which has been forgotten long ago (cf. 7 for references to the original works of Mobius and Bachmann). In general

$$\text{if } \sum_{b_2} f(x \circledast \ell b_2)\, g(b_2) = h(x \circledast \ell) \text{ then } \sum_{b_2} h(x \circledast \ell b_2)\, \bar{g}'(\ell b_2) \tag{37}$$

Now let me state a few interesting identities which can be understood with very little effort using the above matrix approach For the sake of convenience let me introduce some notations.

$$(f * g)(n) = \sum_{d/n} f\left(\frac{n}{d}\right) g(d) = (g * f)(n) \tag{38}$$

$$(f \# g)(x) = \sum_{d \le x} f\left(\frac{x}{d}\right) g(d) \tag{39}$$

$$(f \cdot g)(n) = f(n)\, g(n) \tag{40}$$

$$(f \hat{\#} g)(x) = \sum_{d} f(x, d)\, g(d) \tag{41}$$

The product (40) corresponds to the Madamard product of matrices namely $(F \cdot G)_{ij} = (F_{ij}\, G_{ij})$ Some familiar functions and their properties are as follows

$$\varepsilon(n) = \delta_{n1} \tag{42}$$

$$\varepsilon * f = f * \varepsilon = f \tag{43}$$

$$\nu * \mu = \mu * \nu = \varepsilon \tag{44}$$

$$g_k(n) = n^k \tag{45}$$

$$J_k = g_k * \mu = \mu * g_k \quad \text{(Jordan totient function)} \tag{46}$$

$$J_1 = \phi = g_1 * \mu = \mu * g_1 \quad \text{(Euler function)} \tag{47}$$

Associativity and commutativity of the matrix multiplication representing $*$ gives

$$(f * g) * h = f * (g * h) \tag{48}$$

Thus

$$\phi * \nu = \nu * \phi = g, \tag{49}$$

and

$$\phi * g_k \equiv \sum_{\rho=1}^{n} (g.c.d(\rho, n))^k = (g_1 * \mu) * g_k = g_1 * J_k \tag{50}$$

This identity has important applications in the study of generalized Clifford groups (cf. [2]) (see also [3] for some more interesting aspects of this identity). Buschman [6] has derived the following identity.

$$(f \# g) \# h = f \# (g * h) \tag{51}$$

This is obvious once you interpret the meaning of the operations $\#$ and $*$ in terms of the corresponding matrix operations as detailed above and use the fact of associativity of matrix multiplication. In fact if you take $g = \nu$ and $h = \mu$ it is just the formula (33). Similarly one can have

$$(f \hat{\#} g) \hat{\#} h = f \hat{\#} (g * h) \tag{52}$$

and the Mobius inversion formula (36) results by taking $g = \nu$ and $h = \mu$. Similar simple arguments prove the following

$$g \# \phi = (g \# \mu) \# i \tag{53}$$

$$\begin{aligned} g \# f &= (g \# \mu) \# (\nu * f) \\ &= (g \# \nu) \# (\mu * f) \end{aligned} \tag{54}$$

These lead to the well-known identities

$$L(f,s) \equiv \sum_{n=1}^{\infty} f(n)/n^s \qquad (55)$$

$$= \zeta(s)\, L(\mu * f, s)$$

$$= \frac{1}{\zeta(s)}\, L(\nu * f, s)$$

where $\zeta(s)$ is the Riemann Zetz function

$$g \hat{\#} f = (g \hat{\#} \mu) \hat{\#} (\nu * f) = (g \hat{\#} \nu) \hat{\#} (\mu * f) \qquad (56)$$

Let ℓ denote the logarithm function i.e. $\ell(x) = \log x$

$$((g \# f)\ell) \# h = (g \cdot \ell) \# (f * h) + g \# ((\ell \cdot f) * h) \qquad (57)$$

This is the general form of Tatuzawa-Iseki identity (cf. [6])
The original Tatuza-Iseki identity

$$\text{if } d = (g \# \nu) \cdot \ell \text{ then } d \# \mu = g \cdot \ell + g \# \Lambda \qquad (58)$$

with $\Lambda = (\ell \cdot \nu) * \mu =$ Von-Mangoldt function results by taking $f = \nu$ and $h = \mu$ in (57). The analogues of (57)-(58) for $\hat{\#}$ are

$$((g \hat{\#} f) \cdot \ell) \hat{\#} h = (g \cdot \ell) \hat{\#} (f * h) - g \hat{\#} ((\ell \cdot f) * h) \qquad (59)$$

$$\text{if } d = (g \hat{\#} \nu) \cdot \ell \text{ then } d \hat{\#} \mu = g \cdot \ell - g \hat{\#} \Lambda \qquad (60)$$

qirh aimilar pattern of arguments one can prove the following for $\hat{\#}$ and their analogues for $\#$ given by [6]

$$(h \hat{\#} g) \hat{\#} f = (h \hat{\#} f) \hat{\#} g \tag{61}$$

$$(f \hat{\#} g) \hat{\#} (h \hat{\#} j) = (f \hat{\#} (h \# j)) \hat{\#} g \tag{62}$$

$$f \hat{\#} (g \hat{\#} h) = ((f \hat{\#} \nu) \hat{\#} (g \hat{\#} h)) \hat{\#} \mu \tag{63}$$

$$\text{if } d = (g \hat{\#} \mu) \cdot l \quad \text{then } d \hat{\#} \mu = l \cdot g + g \hat{\#} \Lambda \tag{64}$$

$$((f \hat{\#} g) \cdot l) \hat{\#} h = ((f \cdot l) \hat{\#} g) \hat{\#} h - (f \hat{\#} (g \cdot l)) \hat{\#} h \tag{65}$$

$$((f \hat{\#} g) \cdot l \hat{\#} h = l \cdot (f \hat{\#} (g * h)) + f \hat{\#} ((h \cdot l) * g) \tag{66}$$

Acknowledgement

I am indebeted to Mr.Krishnaswami Alladi for introducing to me the basic concepts in number theory and useful discussions. My sincere thanks are due to Professor Alladi Ramakrishnan and Professor N.R.Ranganathan for kind encouragement.

References

1. Alladi Ramakrishnan, 'L-Matrix Theory or Grammar of Dirac Matrices', (Tata McGraw Hill, Bombay/New Delhi, 1972)

2. R.Jagannathan and N.R.Ranganathan, Reports on Mathematical Physics (Warsaw, Poland), 5, 131 (1974), 7, 229 (1975).

3. Krishnaswami Alladi, Matscience Report 83 (1975) (cf. the last article)

4. D.A.Smith, Lecture notes in Mathematics (Springer Verlag 1971) 251, 205.

5. D.Rearick, Duke Mathematical Journal 35, 767 (1968)

6. R.G.Buschman, Proceedings of American Mathematical Society 25, 307, (1970).

7. L.E.Dickson, History of the Theory of Numbers, Vol.I (Chelsea Pub. Co. New York 1952) 441-445.

ANALOGUES TO THE HARDY-RAMANUJAN THEOREMS*

Krishnaswami Alladi
Mathematics Department, University of California
Los Angeles, California, USA

+*+*+*+

Introduction

1. In this lecture I shall summarise the work done by me in collaboration with P.Erdos and some recent results which shall be the contents of a forthcoming paper. We obtain analogues to the Hardy-Ramanujan theorems on the number of prime factors of an integer, and deduce fruther interesting results on the large prime factors of an integer.

2. The Hardy-Ramanujan Theorems.

It is indeed a little strange that though prime numbers were discussed as far back as Euclid, it is only recently that Hardy-Ramanujan made a discussion of the number of prime factors of an integer. They wanted a mathematical explanation of the phenomenon 'round numbers are rare' where a number may be thought of as round if it is a product of a large number of relatively small prime factors. Thus $2.3.5.7.11.13$ can be called round. Now if we consider $2^5.3^8$ this number has only two distinct prime factors. However if 2 and three are counted according to multiplicity then $2^5.3^8$ is round. In other words a number may be said to be round if it has a large number of prime factors counted either distinctly or multiply (if it ends in zeros when expressed in a large number of bases, or ends in a large number of zeros when expressed in possibly few bases). To make our statements precise we need

* Lecture delivered at the Matscience Symposium on Matrix Theory, Computational methods and Number theory September 8, Mysore

a measure of roundness of a number. Thus the two functions of Hardy and Ramanujan are the following:

Let an integer n be expressed as a product of distinct primes $n = \prod_{i=1}^{r} p_i^{\alpha_i}$. Set

$$\Omega(n) = \sum_{i=1}^{r} \alpha_i \; ; \quad \omega(n) = \sum_{i=1}^{r} 1 = r$$

It is obvious that both $\Omega(n)$ and $\omega(n)$ fluctuate in size. Nevertheless their averages are well behaved, and nearly the same ! This is expressed in

THEOREM 1. a) $\sum_{1 \le n \le x}' \Omega(n) = x \log\log x + B_1 x + o(x)$

b) $\sum_{1 \le n \le x}' \omega(n) = x \log\log x + B_2 x + o(x)$

$\left(\sum_{1 \le n \le x}' \log\log n = x \log\log x + O(x) \right)$.

In other words the average of <u>both</u> these functions behaves like loglogn. An average is essentially influenced by two things. Firstly by the values a function takes most often, and secondly by the abnormally large values of a function. We would like to know what exactly happens in this case. This is expressed in the second theorem of Hardy-Ramanujan given below, namely, that both the functions are nearly loglogn almost always.

THEOREM 2. Choose $\varepsilon > 0$ arbitrary. Denote by

$$\psi_\varepsilon(x) = \#\text{ of } n \le x \text{ for which } 1-\varepsilon < \frac{\Omega(n)}{\log\log n} < 1+\varepsilon$$

Then

$$\lim_{x\to\infty} \frac{\psi_\varepsilon(x)}{x} = 1 \quad \forall \varepsilon > 0$$

The above statement holds if $\Omega(n)$ is replaced by $\omega(n)$. From theorem 2 it is clear that $\Omega(n)$ and $\omega(n)$ are nearly the same almost always. Thus the two measures of roundness coincide most often. In otherwords, most prime factors of a number occur single. Also, most numbers have nearly loglogn prime factors according to both measures, a number that is large enough to make the average value of α_i behave like unity, still small compared to the size of the number, which explains the phenomenon that round members are rare. For a proof of these two theorems see

3. Analogues and Extensions:

A few years agao the author studied the sum of the prime factors of an integer which arose naturally out of the partitions of integers into primes (see [1]). It must again be mentioned that this function (denoted by $A(n)$ below) has not been given much attention. The function becomes extremely interesting when looked upon in the light of the Hardy-Ramanujan theorems and for this we refer to a paper by the author and Erdos [2], which includes the other pertinent references regarding $A(n)$.

As before there are two ways of summing the prime factors of an integer. Denote by

$$A(n) = \sum_{i=1}^{r} \alpha_i p_i \; ; \quad A^*(n) = \sum_{i=1}^{r} p_i$$

where $n = \prod_{i=1}^{r} p_i^{\alpha_i}$. It is obvious that the functions fluctuate in size wildly. However their average is well behaved and nearly the same.

THEOREM 3.

$$\text{a)} \sum_{1 \le n \le x}{}' A(n) = \frac{\pi^2 x^2}{12 \log x} + O\left(\frac{x^2}{\log^2 x}\right)$$

$$\text{b)} \sum_{1 \le n \le x}{}' A^*(n) = \frac{\pi^2 x^2}{12 \log x} + O\left(\frac{x^2}{\log^2 x}\right)$$

$$\left(\sum_{1 \le n \le x}{}' \frac{\pi^2 n}{6 \log n} = \frac{\pi^2 x^2}{12 \log x} + O\left(\frac{x^2}{\log^2 x}\right) \right)$$

Theorem 3 says that both the functions have the same average, that is $\pi^2 n / 6 \log n$. The next question is whether this average is influenced by the values taken most often or by the abnormally large values of the function.

It is a theorem of Erdos that the largest prime factor of n, is never nearly equal to well behaved function in the sense of theorem 2. Also it is reasonable to guess that in the sum A(n) the largest prime factor of n ($P_1(n)$ say) dominates the others

somuch so that $\Lambda(n)$ and $P_1(n)$ are nearly the same most often. Thus it would be hopeless to try for an analogue to theorem 3. So the average here is influenced by the abnormally large values of $\Lambda(n)$ and $\Lambda^*(n)$.

One can show that

$$\sum_{1\le n\le x}{}' \{\Lambda(n) - \Lambda^*(n)\} = x \log\log x + O(x) \qquad (1)$$

Observe that $\Lambda(n) \geqslant \log n$ so that

$$\sum_{1\le n\le x}{}' \left\{1 - \frac{\Lambda^*(n)}{\Lambda(n)}\right\} \le \sum_{1\le n\le x}{}' \frac{\Lambda(n) - \Lambda^*(n)}{\log n} = O\left(\frac{x \log\log x}{\log x}\right) = o(x) \qquad (2)$$

Clearly $1 - \frac{\Lambda^*(n)}{\Lambda(n)} \geqslant 0$. Now let

$$x - \psi_\varepsilon(x) = \#\, n \le x \text{ for which } 1 - \frac{\Lambda^*(n)}{\Lambda(n)} > \varepsilon$$

$$\text{or } \psi_\varepsilon(x) = \#\, n \le x \text{ for which } 1 < \frac{\Lambda^*(n)}{\Lambda(n)} < 1+\varepsilon$$

Then the left side consisting only of positive quantities is $\geqslant \{x - \psi_\varepsilon(x)\} - \varepsilon$. Thus

$$\varepsilon\left(x - \psi_\varepsilon(x)\right) = O\left(\frac{x \log\log x}{\log x}\right) \qquad (3)$$

from (2). Now since $\varepsilon > 0$, (3) implies that $\psi_\varepsilon(x)/x \to 1$ as $x \to \infty$ for every $\varepsilon > 0$. In otherwords $\Lambda(n)$ and $\Lambda^*(n)$ are nearly the same almost always. Thus we have a partial analogue to Theorem 2.

THEOREM 4. The functions $\Lambda(n)$ and $\Lambda^*(n)$ are nearly the same almost always.

We say that two functions and are nearly the same almost always if for each $\varepsilon > 0$, $\lim_{x\to\infty} \frac{\psi_\varepsilon(x)}{x} = 1$ where

$$\psi_\varepsilon(x) = \#\, n \le x \text{ such that } 1-\varepsilon < \frac{f(n)}{g(n)} < 1+\varepsilon.$$

Now let us look at $\Lambda(n)$ and $\Lambda^*(n)$ in the light of the Hardy-Ramanujan theorems and the function $\Lambda(n)$. According to Theorems 1 and 2 most prime factors occur single. So it is natural to believe that the large prime factors occur single with greater frequency. We have already mentioned before that we expect $P_1(n)$ to dominate over the other primes in both the sums $\Lambda(n)$ and $\Lambda^*(n)$. Now if we expend this idea we can expect $\Lambda(n) - P_1(n) - P_2(n) - \cdots - P_{m-1}(n)$ and $\Lambda^*(n) - P_1(n) - P_2(n) - \cdots - P_{m-1}(n)$ to behave like $P_m(n)$ where $P_k(n)$ represents the k^{th} largest prime factor of n, which is set equal to zero the moment we run out of primes. Since these functions are wildly fluctuating it is interesting that their averages exist. So one of the main theorems is

THEOREM 5. Let m be a fixed integer > 0. Then

$$\sum_{1\le n\le x} \{\Lambda(n) - P_1(n) - \cdots - P_{m-1}(n)\} \sim \sum_{1\le n\le x} \{\Lambda^*(n) - P_1(n) - \cdots - P_{m-1}(n)\}$$

$$\sim \sum_{1\le n\le x} P_m(n) \sim \frac{c_m x^{1+\frac{1}{m}}}{(\log x)^m}$$

where c_m is a positive constant which is a rational multiple of $\zeta(1+\frac{1}{m})$ where ζ is the Riemann-Zeta function.

Theorem 5 is an analogue and extension of Theorem 1 and obviously says much more about large prime factors. Now we want an analogue to Theorem 2. It is certainly hopeless to find an exact analogue for by Erdos theorem one cannot find a well behaved function to be nearly equal to these quantities almost always. Nevertheless we can prove a theorem similar to theorem 4 This another main theorem is

THEOREM 6. The functions $\Lambda(n) - P_1(n) - P_2(n) - \ldots - P_{m-1}(n)$, $P_m(n)$ and $\Lambda^*(n) - P_1(n) - P_2(n) - \ldots - P_{m-1}(n)$ are nearly the same almost always.

It is quite clear that theorem 5 and 6 are extensions of Theorems 3 and 4 which are analogues to Theorems 1 and 2. A very interesting remark accompanies theorems 5 and 6. Consider a number $2^3 . 8^2 . 5^4 = n$. Clearly $P_1(n) = 5$. As regards $P_2(n)$ we may define it as either 3 or 5 depending whether we choose the strict inequality $(<)$ or the weak one $(\leq)$ for our definition Now there are 2^{k-1} possible definitions for the k^{th} largest prime factor. Remarkably however we have

THEOREM 7. Theorems 5 and 6 hold no matter which of the 2^{k-1} definitions for $P_k(n)$ we choose for $k = 1,2,\ldots,m$.

Theorem 6 holds for any one of these definitions because

THEOREM 8. For every fixed k, $P_k(n)$ occurs square-free for almost all n, where $P_k(n)$ is the k^{th} largest prime factor defined by strict inequalities.

As regards Theorem 5 it has been proved in the paper of Alladi and Erdos [2] for $P_k(n)$ defined only by weak inequalities. On observing the asymptotic terms in the proof one deduces that the same estimates hold if weak inequalities are replaced by strict ones.

4. Methods of Proof:

Theorems 3, 4 and 5 may be found in [2] . The proof of Theorems 5 is very long and rests on tricky estimates of sums involving primes. Since it is too cumbersome for insertion here we do not go through it. Theorem 6 is more recent and will appear with other results in a forthcoming paper. We shall sketch the proof of a particular case of Theorem 6 i.e. m = 1. The proof for the general case involves the same method but is a little more complicated.

Consider Theorem 6 for m = 1. We need to show that $\Lambda(n)$, $\Lambda^*(n)$ and $P_1(n)$ are nearly the same almost always, Theorem 4 which showed that $\Lambda(n)$ and $\Lambda^*(n)$ are nearly the same was deduced directly from (1), (2) and (3). Now by Theorem 5

$$\sum_{1\le n\le x}{}' \{\Lambda(n) - P_1(n)\} \sim \sum_{1\le n\le x}{}' \{\Lambda^*(n) - P_1(n)\} \sim \sum_{1\le n\le x}{}' P_2(n) \sim c_2 \frac{x^{1+\frac{1}{2}}}{(\log x)^2}$$

So the method in (1) and (2) which involves division by $A(n)$ and the bound $A(n) \geq \log n$ does not work here.

This however can still be proved using interesting ideas of duality between large and small prime factors of integers.

Define a function f by

$$f(n) = \begin{cases} p & \text{if } n = p^m \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Clearly

$$A(n) = \sum_{d|n}{}' f(d)$$

so that if μ represents the Moebius function we have

$$f(n) = \sum_{d|n}{}' \mu(d) A\left(\frac{n}{d}\right) = -\sum_{d|n}{}' \mu(d) A(d)$$

and so

$$\sum_{d|n}{}' \mu(d) A^*(d) = \sum_{d|n}{}' \mu(d) A(d) = -f(n) \tag{5}$$

Now let us consider the sum

$$\sum_{d|n}{}' \frac{\mu(d) A^*(d)}{P_1(d)}$$

where we define

$$\mu(1) \frac{A^*(1)}{P_1(1)} = \frac{1.0}{0} = 1.$$

Now let $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ where the p_i are distinct primes with $p_1 < p_2 < \dots < p_r$. Now we may split the above sum as

$$\sum_{d|n}{}' \frac{\mu(d) A^*(d)}{P_1(d)} = 1 + \sum_{j=1}^{r}{}' \frac{1}{p_j} \sum_{\substack{d|n \\ P_1(d)=p_j}}{}' \mu(d) A^*(d) \qquad (6)$$

Now $P_1(d) = p_j$ if and only if

$$d = p_j e \quad \text{where} \quad e \Big| p_1^{\alpha_1} \dots p_{j-1}^{\alpha_{j-1}} \cdot p_j^{\alpha_j - 1}$$

However if $p_j^2 \mid d$, $\mu(d) = 0$. So we may assume that $(e, p_j) = 1$ or that $e \mid p_1^{\alpha_1} \dots p_{j-1}^{\alpha_{j-1}} = d'$ say. So in view of this observation we rewrite (6) as

$$\sum_{d|n}{}' \frac{\mu(d) A^*(d)}{P_1(d)} = 1 + \sum_{j=1}^{r}{}' \frac{1}{p_j} \sum_{e \mid p_1^{\alpha_1} \dots p_{j-1}^{\alpha_{j-1}}}{}' \mu(p_j e) A^*(p_j e)$$

$$= 1 + \sum_{j=1}^{r}{}' \frac{\mu(p_j)}{p_j} \left\{ \sum_{e \mid p_1^{\alpha_1} \dots p_{j-1}^{\alpha_{j-1}}}{}' \mu(e) A^*(e) + A^*(p_j) \sum_{e \mid p_1^{\alpha_1} \dots p_{j-1}^{\alpha_{j-1}}}{}' \mu(e) \right\}$$

Now if $j > 2$ then because of (5) we infer that both sums inside the brackets vanish. If $j = 2$ then the latter sum vanishes and (5) gives p_1/p_2. If $j = 1$ we get $\frac{-p_1}{p_1}$. Thus if $r \geq 2$ we get

$$\sum_{d|n} \frac{\mu(d) A^*(d)}{P_1(d)} = 1 + \frac{p_1}{p_2} - \frac{p_1}{p_1} = \frac{p_1}{p_2}$$

If $r = 0$ we get the sum $= 1$. If $r = 1$ we get the sum $= 0$. Thus

$$\sum_{d|n} \frac{\mu(d) A^*(d)}{P_1(d)} = \frac{p_1^*(n)}{p_2^*(n)} \tag{7}$$

where $p_1^*(n)$ and $p_2^*(n)$ are defined as follows:

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r} \; ; \; p_1 < p_2 < \cdots < p_r$$

$r \geq 2$: $p_1^*(n) = p_1$, $p_2^*(n) = p_2$

$r = 1$ $p_1^*(n) = 0$, $p_2^*(n) = p_1$

$r = 0$ $p_1^*(n) = 0 = p_2^*(n)$ so that $\frac{0}{0} = 1$.

(The summation process in (6) leads to a very interesting duality notion between large and small prime factors of integers which we shall treat in § 5).

Now from (7) we infer from Moebius inversion that

$$\frac{\mu(n)\,A^*(n)}{P_1(n)} = \sum_{d|n}{}' \mu\left(\frac{n}{d}\right) \frac{P_1^*(d)}{P_2^*(d)} \quad . \tag{8}$$

Let us assume that n is square free. Then (8) leads to

$$\frac{A^*(n)}{P_1(n)} = \sum_{d|n}{}' \mu(d) \frac{P_1^*(d)}{P_2^*(d)} \tag{9}$$

Now if $\gamma(n)$ represents the largest square free divisor of n then observe that

$$\sum_{d|n}{}' \mu(d) \frac{P_1^*(d)}{P_2^*(d)} = \sum_{d|\gamma(n)}{}' \mu(d) \frac{P_1^*(d)}{P_2^*(d)} = \frac{A^*(\gamma(n))}{P_1(\gamma(n))} = \frac{A^*(n)}{P_1(n)}$$

so that (9) holds for all n.

Now consider

$$\sum_{1\le n\le x}{}' \frac{A^*(n)}{P_1(n)} = \sum_{1\le n\le x}{}' \sum_{d|n}{}' \mu(d) \frac{P_1^*(d)}{P_2^*(d)}$$

$$= \sum_{1\le d\le x}{}' \mu(d) \frac{P_1^*(d)}{P_2^*(d)} \left[\frac{x}{d}\right] = [x] + \sum_{2\le d\le x}{}' \mu(d) \frac{P_1^*(d)}{P_2^*(d)} \left[\frac{x}{d}\right]$$

(10)

We may write

$$\sum_{2\le d\le x}{}' \mu(d)\frac{p_1^*(d)}{p_2^*(d)}\left[\frac{x}{d}\right] = \sum_{p_1,p_2}{}' \sum_{2\le d\le x}{}' \mu(d)\frac{p_1^*(d)}{p_2^*(d)}\left[\frac{x}{d}\right]$$

$$p_1^*(d) = p_1 \quad (p_1 < p_2 \text{ are primes})$$
$$p_2^*(d) = p_2 \qquad (11)$$

Clearly

$$\sum_{\substack{2\le d\le x \\ p_1^*(d)=p_1 \\ p_2^*(d)=p_2}}{}' \mu(d)\frac{p_1^*(d)}{p_2^*(d)}\left[\frac{x}{d}\right] = \frac{p_1}{p_2}\sum_{\substack{e\le \frac{x}{p_1p_2} \\ (e,N)=1}}{}' \mu(p^\alpha q^\beta e)\left[\frac{x}{p^\alpha q^\beta e}\right]$$

$$= \frac{p_1}{p_2}\sum_{\substack{e\le \frac{x}{p_1p_2} \\ (e,N)=1}}{}' \mu(p_1p_2e)\cdot\left[\frac{x}{p_1p_2e}\right] \qquad (12)$$

where $N = \prod_{p\text{-prime}\le p_2} p$. Set $y = x/p_1p_2$. We can write (12) as

$$\frac{p_1}{p_2}\sum_{\substack{e\le y \\ (e,N)=1}}{}' \mu(e)\left[\frac{y}{e}\right] = \frac{p_1}{p_2}\sum_{1\le n\le y}{}' \sum_{\substack{e\mid n \\ (e,N)=1}}{}' \mu(e)$$

$$= \frac{p_1}{p_2}\sum_{\substack{1\le n\le y \\ P_1(n)\le P_1(N)=p_2}}{}' 1 \qquad (13)$$

Now if $\psi(y,p)$ represents the number of $n \leq y$ with $P_1(n) \leq p$ then (12), (13) and (11) imply

$$\sum_{2\leq d\leq x}{}' \mu(d)\frac{p_1^*(d)}{p_2^*(d)}\left[\frac{x}{d}\right] = \sum_{p_1,p_2}{}' \frac{p_1}{p_2}\psi(y,p_2)$$

$$= \sum_{\substack{p_1,p_2 \\ p_2\leq \log x}}{}' \frac{p_1}{p_2}\psi(y,p_2) + \sum_{\substack{p_1,p_2 \\ p_2>\log x}}{}' \frac{p_1}{p_2}\psi(y,p_2)$$

$$= S_1 + S_2 \quad \text{respectively.} \tag{14}$$

Now

$$S_2 = \sum_{\substack{p_1,p_2 \\ p_2>\log x}}{}' \frac{p_1}{p_2}\psi(y,p_2) \leq \sum_{\substack{p_1,p_2 \\ p_2>\log x}}{}' \frac{p_1}{p_2}\frac{x}{p_1p_2} = x\sum_{\substack{p_1,p_2 \\ p_2>\log x}}{}' \frac{1}{p_2^2}$$

$$= x\sum_{p_2>\log x}{}' \frac{1}{p_2^2}\sum_{p_1<p_2}{}' 1 = O\left(x\sum_{p_2>\log x}{}' \frac{1}{p_2\log p_2}\right)$$

$$= o(x) \tag{15}$$

because the series

$$\sum_{q\text{-prime}}{}' \frac{1}{q\log q} < \infty$$

Now if $p = y^{1/u}$ then by a result of de Bruijn [3] we have

$$\psi(y,p) < \frac{y}{e^{u\log u}}$$

But in S_1 $y = \frac{x}{p_1p_2} > \frac{x}{\log^2 x}$. Also $\log x > p_2 = y^{\frac{1}{u}}$, so that $u > \frac{\log x}{\log\log x}$. Thus

$$S_1 = O\left(\sum_{p_1,p_2}{}' \frac{x}{\log^2 x\, e^{u\log u}}\right) = o(x)$$

(16)

So by (15) and (16) we infer

$$\sum_{2\le d\le x}{}' \mu(d)\,\frac{p_1^*(d)}{p_2^*(d)}\left[\frac{x}{d}\right] = o(x)$$

Thus by (10) we deduce that

$$\sum_{1\le n\le x}{}' \frac{A^*(n)}{P_1(n)} \sim x$$

We rewrite this as

$$\sum_{1\le n\le x}{}' \left\{1 - \frac{A^*(n)}{P_1(n)}\right\} = o(x) \qquad (17)$$

Clearly $1-\frac{A^*(n)}{P_1(n)} \le 0$. Now denote by

$$\psi_\varepsilon(x) = \#\, n\le x : 1 \le \frac{A^*(n)}{P_1(n)} < 1+\varepsilon$$

or

$$x - \psi_\varepsilon(x) = \#\, n\le x : 1 - \frac{A^*(n)}{P_1(n)} < -\varepsilon$$

Then from (17) we have

$$\{x\,\psi_\varepsilon(x)\}\varepsilon = o(x) \quad \forall\, \varepsilon > 0$$

or

$$\lim_{x\to\infty} \frac{\psi_\varepsilon(x)}{x} = 1 \quad \forall\, \varepsilon > 0$$

which proves that $\Lambda^*(n)$ is nearly $P_1(n)$ almost always. Now from Theorem 4 we infer that for $m = 1$ Theorem 6 is true.

5. Duality between large and small prime factors.

The summation involved in (6) leads to a new notion of duality between large and small prime factors of an integer. Let g and f be arithmetical functions. Now consider a summation of the type

$$\sum_{d|n} g(d)\, f(P_1(d)) \tag{18}$$

This may be written as (if $n = \prod_{i=1}^{r} p_i^{\alpha_i}$, $p_1 < p_2 < \cdots < p_r$

$$g(1)f(0) + \sum_{j=1}^{r} f(p_j) \sum_{\substack{d|n \\ P_1(d) = p_j}} g(d)$$

$$= g(1)f(0) + \sum_{j=1}^{r} f(p_j) \sum_{e|d'} g(p_j e)$$

(19)

where $d' = p_1^{\alpha_1} \ldots\ldots p_r^{\alpha_r}$. Thus if g is an arithmetic function such that

$$\sum_{d|n}' g(d)$$

vanishes over a set say of density one, then the sum in (18) has a simpler expression. The duality arises if we reverse the ordering and consider instead a sum

$$\sum_{d|n}' g(d) f(p_1(d))$$

where p_1 is the smallest prime factor of d. If further g is assumed to be multiplicative or additive then the expression becomes simpler. In our Theorem 6 the identity involved $g = \mu.\Lambda^*$ and $f(n) = \frac{1}{n}$. Note that in our case

$$\sum_{d|n}' g(d) = \begin{cases} 0 & \text{if } n \neq p^m \\ -p & \text{otherwise.} \end{cases}$$

This led to the identity

$$\sum_{d|n}' g(d) f(P_1(d)) = \frac{p_1^*(d)}{p_2^*(d)}$$

which later helped us prove theorem 6. So one obvious function to choose is $g = \mu$. In this case an analysis similar to (6) yields the following two strikingly simple and elegant identities which brings out the duality between large and small prime factors of an integer. The proofs of these identities use the principle underlying (19) . It is indeed remarkable that if f is any function defined on the integers then

$$\sum_{d|n} \mu(d)\ f(P_1(d)) = f(o) - f(p_1(n)) \qquad (20)$$

and

$$\sum_{d|n} \mu(d)\ f(p_1(d)) = f(o) - f(P_1(n)) \qquad (21)$$

In particular setting $f(o) = 0$ we get

$$\sum_{d|n}' \mu(d)\ f(P_1(d)) = - f(p_1(n)) \qquad (22)$$

$$\sum_{d|n} \mu(d)\ f(p_1(d)) = - f(P_1(n)) \qquad (23)$$

Now subtracting (21) from (20) we get

$$\sum_{d|n}' \mu(d) \left\{ f(P_1(d)) - f(p_1(d)) \right\} = f(P_1(n)) - f(p_1(n)) \qquad (24)$$

Now if we set

$$g(n) = f(P_1(n)) - f(p_1(n)) \qquad (25)$$

then (24) takes the shape

$$\sum_{d|n}' \mu(d)\ g(d) = g(n) \qquad (26)$$

An interesting problem would be to determine that given g satisfying (26) under what conditions can it be written in the form (25) ? Similarly if $h(n) = f(P_1(n)) + f(p_1(n))$, then

$$\sum_{d|n}' \mu(d)\ h(d) = -h(n) \qquad (f(o) = 0)$$

The converse question here again can be raised. These are elegant duality relations connecting $P_1(n)$ to $p_k(n)$ and $p_1(n)$ to $P_k(n)$ where P_k and p_k are the k^{th} largest and k^{th} smallest prime factors of n respectively defined by strict inequalities only. Let f be an arithmetic function with $f(o) = 0$. Then

$$\sum_{d|n}{}' \mu(d)\, f(P_k(d)) = (-1)^k \binom{\omega(n)-1}{k-1} f(p_1(n)) \qquad (27)$$

$$\sum_{d|n}{}' \mu(d)\, f(p_k(d)) = (-1)^k \binom{\omega(n)-1}{k-1} f(P_1(n)) \qquad (28)$$

where $\omega(n)$ is the function discussed in Theorem 1. Identities similar to (27) and (28) are used in the proof of Theorem 6 for the general case of arbitrary . This and other results relating to a deeper analysis of duality will be the contents of a forthcoming paper.

REFERENCES.

(1) K.Alladi, A New Logarithmic function, Srinivasa Ramanujan Commemoration Volume, Madras, India 1974, part 2, pp.

(2) K.Alladi and P.Erods, On an Additive Arithmetic function, Pacific Journal of Mathematics, To appear.

(3) N.J. de Bruijn, On the number of integers $\leq x$ free of prime factors $> y$. Indag. Math (1950).

(4) G.H.Hardy and E.M.Wright, An Introduction to theory of Numbers, Clarendon, Oxford.

NSS/13.9.1976

SOME COMPUTATIONAL METHODS IN PHYSICS

K.Srinivasa Rao
MATSCIENCE, The Institute of Mathematical Sciences,
Madras-20. (INDIA)

***

The advent of high-speed digital computers, approximately twenty-five years ago, has made available to scientists an altogether new medium using which it is now possible to solve, numerically, mathematical problems, as efficiently as possible. But, it should be noted that, however large the computing system may be, it is still finite in comparison with the number of atoms in a liquid or the number of stars in a galaxy and one has essentially to formulate finite numerical models of the few-body and many-body problems that one encounters frequently. The ultimate goal of a scientist is to develop a theory based on as small a number of principles as possible to describe a system mathematically as simply, as exactly, and as completely as possible, such that the theory simulates the experimental data.

Computer problems arise in almost all branches of physics, classical or quantum mechanical , which demand considerable amount of computational analysis. Frequently, one encounters nonlinear (partial differential) equations, whose solution is seldom feasible without the use of the computer. The advent of manipulation of symbols or algebraic expressions by a computer, has enlarged the domain of computational physics beyond the realms of numerical calculations. Thus the field of computational physics open before

me is so vast, that it is impossible for me to do justice to this rapidly growing field in just two forty-five minute talks. So, I restrict myself to the theme of the Seminar by talking about the use of the method of finite differences to reduce second order differential equations to a system of simultaneous equations which can be solved by using matrix methods.

Solution of problems which merit numerical computation on a digital computer involve the following three stages:

(a) formulation and choice of method

(b) programming and coding for a computer

(c) 'debugging' by means of test runs on a computer

Hence, we will be concerned with the first aspect, adequate attention to which is as important, to success in computation, as efficiently, programming and coding the method for use on a computer.

I wish to mention that efficient programmes exist for solving linear equations, finding eigenvalues and eigenvectors of matrices solving certain classes of differential equations, performing certain kinds of numerical quadrature, generating random numbers, curve fitting procedures, etc. These programmes which are sought after, by scientists in many fields have been tested and computer centres without exception have useful programme libraries. In passing, it should be mentioned that nuclear physicists and crystallographers, to mention two groups of scientists, who resort to large scale computations have their own collection of programmes developed by themselves or by other experts.

In the field of numerical matrix algebra, two problems of interest to computational physicists are: the solution of the matrix equation and the determination of the eigenvalues and eigenvectors of particular matrices. In the former, we require to solve the set of simultaneous equations which relate the unknown variables $u_1, \ldots, u_n$ to the known variables $w_1, \ldots, w_n$:

$$\begin{aligned} a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n &= w_1, \\ a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n &= w_2, \\ \vdots \qquad\quad \vdots \qquad\quad & \qquad\quad \vdots \\ a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n &= w_m \end{aligned} \tag{1}$$

These can be written in matrix notation as:

$$\underline{A}\,\underline{u} = \underline{w} \tag{2}$$

where A is a m x n matrix, u is a n-component column vector and w is a m-component column vector. When the number of rows is equal to the number of columns of the matrix A, i.e. m = n, then, we can obtain the solution:

$$\underline{u} = \underline{A}^{-1}\underline{w} \tag{3}$$

where A^{-1}, the inverse of the matrix A, exists if A is a non-singular matrix (i.e. a matrix with determinant not equal to zero). By Cramer's rule, the inverse of a matrix is given by:

$$(\underline{A}^{-1})_{ij} = (-1)^{i+j}\frac{|A|^{ji}}{|A|} \tag{4}$$

where $|A|$ is the determinant of $\underline{A}$ and $|A|^{ji}$ is the determinant of the (n-1) x (n-1) matrix formed by striking out the i^{th} row and jth column of the matrix $\underline{A}$. The determinant itself is defined recursively as:

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} |A|^{ij} a_{ij} \qquad (5)$$

Clearly, the calculation of a single determinant of an n x n matrix alone involves n! multiplications (note: 10! = 36,28,800). The matrices which occur in certain problems, to be mentioned in due course, have large dimensions (n), at least 100 and as large, as 100,000. Obviously, the direct application of Cramer's rule for these problems is out of question.

In what follows, we develop finite difference calculus and show how it is useful in solving a class of
/Boundary-value problems of frequent occurrence in physics.

Consider an independent continuous variable x , which lies in the domain:

$$X = [X_1, X_2]$$

$$X_1 \leq x \leq X_2 \qquad (6)$$

The continuouum is replaced by a mesh, or lattice of points, by dividing the domain of X into a set of N-1 elements, Δx_n as shown in figure below

Fig 1.

A vector $\{x_n\}$ of finite dimension N can then be constructed by defining the continuous variable x at the n points. Thus,

$$x_n = X_1 + \sum_{\nu=1}^{n-1} \Delta x_\nu \tag{7}$$

If f(x) is a dependent function in the continuum, the function f can be approximated by defining a corresponding vector $\{f_n\}$, on the independent variable mesh $\{x_n\}$:

$$f_n = f(x_n) \tag{8}$$

Obviously, since the function f is defined originally everywhere on the continuous variable x, the representation $\{f_n\}$ is an incomplete description of f(x). Nevertheless, the function f may be approximated from $\{f_n\}$ at any point x'.

$$x_n \leq x' \leq x_{n+1} \tag{9}$$

by interpolation from the vector components f_n and f_{n+1} between adjacent points. If

$$\epsilon = \frac{x' - x_n}{x_{n+1} - x_n} = \frac{x' - x_n}{\Delta x_n} \tag{10}$$

then, with first-order interpolation:

$$f^* = \epsilon f_{n+1} + (1 - \epsilon) f_n \tag{11}$$

and, in this sense, f^* approximates f. Obviously, this approximation is a good description for a slowly-varying continuous function f.

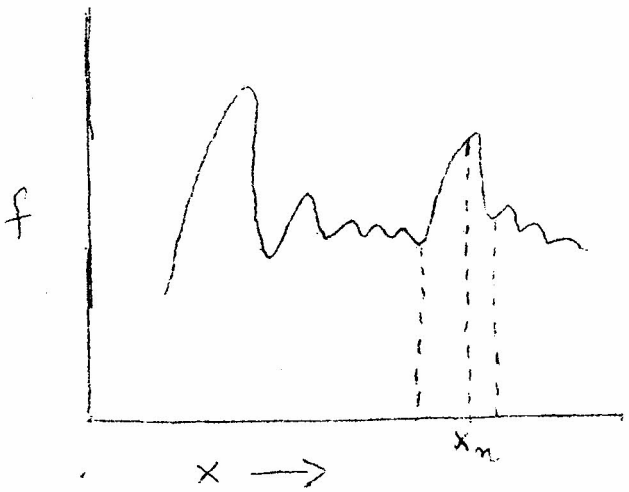


Fig.2(a)

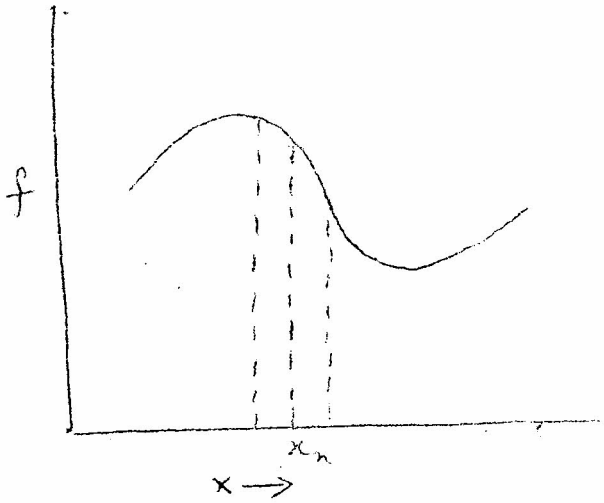


Fig.2(b)

If f changes rapidly over the element Δx_n, then f^* is a 'poor' approximation to f. Clearly, we cannot describe wave lengths less than Δx_n and the more points there are in the interval, the better is the representation $\{f_n\}$ of f. The function shown in Fig.2(a) is more rapidly varying and so more points are needed in its representation.

Having considered the representation of a continuous function on a discrete mesh, we shall now analyse approximations to the differential derivatives of functions. A derivative of a function yields information about the local variation of the function in space and consequently a difference derivative couples neighbouring points on the mesh. The obvious approximation to the first

derivative df/dx at the point n on the mesh, $1 < n < N$ is

$$\Delta'_x f_n = \frac{f_{n+1} - f_{n-1}}{2\Delta} \quad (12)$$

where Δ is the mesh step length. This is a good approximation to the derivative df/dx, if f does not change very rapidly over Δ. To estimate the extent of the approximation Δ'_x for $\frac{d}{dx}$, let $u = g e^{ikx}$

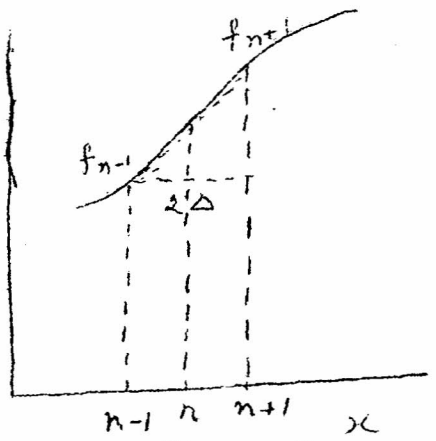


Figure 3

then $\frac{du}{dx} = ikge^{ikx} = iku$.

Operating the difference derivative on the Fourier mode leads to:

$$\Delta'_x u = \frac{1}{2\Delta}[ge^{ikx_{n+1}} - ge^{ikx_{n-1}}]$$

$$= \frac{g}{2\Delta}[e^{ik(x_n+\Delta)} - e^{ik(x_n-\Delta)}] \quad \text{since } x_{n+1} = x_n+\Delta,\ x_{n-1} = x_n-\Delta$$

$$= \frac{g}{\Delta} e^{ikx_n} \frac{1}{i}[e^{ik\Delta} - e^{-ik\Delta}]$$

$$= \frac{iu}{\Delta} \sin k\Delta \sim iku \text{ for small } k\Delta \quad (13)$$

Hence the difference formula is a good approximation for the first derivative when the wave number k is small (the wavelength $\frac{2\pi}{k}$ is large). The longer the wavelength, the better is the approximation. More exactly,

$$\Delta'_x u = \frac{iu}{\Delta}\left(k\Delta - \frac{(k\Delta)^3}{6} + O(k^5\Delta^5)\right)$$

$$= iuk\left(1 - \frac{k^2\Delta^2}{6} + O(k^4\Delta^4)\right)$$

$$\therefore \Delta'_x \equiv \left(1 - \frac{k^2\Delta^2}{6} + O(k^4\Delta^4)\right)\frac{d}{dx} \quad (14)$$

We say the difference operator is second-order accurate in or 'space-centred'.

For the second derivative, clearly a good and simple approximation to $\frac{d^2f}{dx^2}$ on the mesh $1<n<N$ is:

$$\Delta'_x f = \frac{f_{n+\frac{1}{2}} - f_{n-\frac{1}{2}}}{\Delta}$$

$$\Delta''_x f = \frac{f_{n+1} - f_n - (f_n - f_{n-1})}{\Delta^2}$$

$$= \frac{f_{n+1} - 2f_n + f_{n-1}}{\Delta^2} \qquad (15)$$

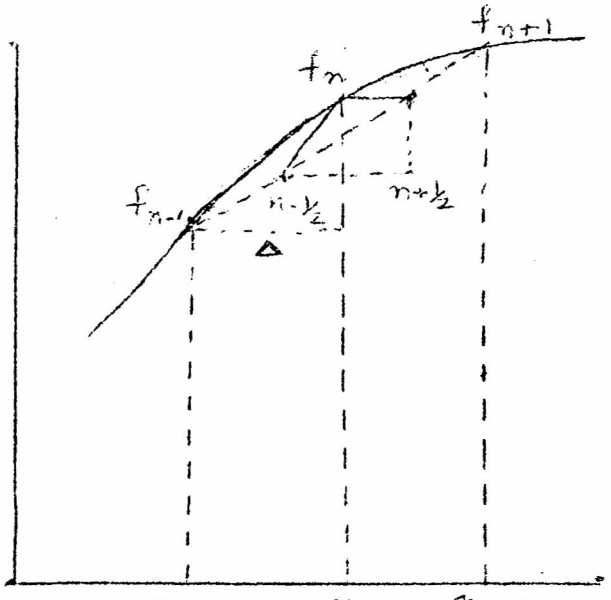


Figure 4

As before, we can establish the accuracy of this approximation by considering the effect of Δ''_x operating on a Fourier mode of wavelength $\frac{2\pi}{k}$. For the differential operator

$$\frac{d^2}{dx^2} = \frac{d^2}{dx^2}(g e^{ikx}) = -gk^2 e^{ikx} = -k^2 u$$

For the second order difference operator, using the above algorithm

$$\Delta''_x u_n = \frac{g}{\Delta^2}\left[e^{ik(x_n+\Delta)} - 2e^{ikx_n} + e^{ik(x_n-\Delta)}\right]$$

$$= \frac{g}{\Delta^2} e^{ikx_n}\left[\tfrac{1}{2}e^{ik\Delta} + \tfrac{1}{2}e^{-ik\Delta} - 1\right]$$

$$= \frac{2u}{\Delta^2}(\cos k\Delta - 1) \sim \frac{2u}{\Delta^2}\left[1 - \frac{k^2\Delta^2}{2} - 1\right]$$

$$= -uk^2 \text{ for small } k\Delta .$$

$$\equiv \frac{2u}{\Delta^2}\left[1 - \frac{k^2\Delta^2}{2} + \frac{k^4\Delta^4}{24} - 1 + O(k^6\Delta^6)\right]$$

$$= -k^2u\left(1 - \frac{k^2\Delta^2}{12} + O(k^4\Delta^4)\right)$$

$$\therefore \Delta_x'' \equiv \left(1 - \frac{k^2\Delta^2}{12} + O(k^4\Delta^4)\right)\frac{d^2}{dx^2} \qquad (16)$$

These approximations to he first and second derivatives are commonly employed in simulating differential equations of interest in physics. The same method of analysis can be employed for higher-order derivatives.

Let us consider the form of matrices which arise from the finite difference calculus. As a simple but important example, let us examine the solutions of a one-dimensional Poisson's equation:

$$\frac{d^2\Phi(x)}{dx^2} = -\rho(x) \qquad (17)$$

where $\rho(x)$ is known as the source function, while $\Phi(x)$ is the unknown potential function. On an equally spaced difference mesh (space step Δ) $1 \leq n \leq N$, the function $\rho(x)$ is replaced by a vector $\{\rho_n\}$ defined at mesh points x_n. Approximating the second-order differential operator d^2/dx^2 by the difference operator Δ_x'', we get for the differential equation, the following set of equations:

$$\Phi_{n+1} - 2\Phi_n + \Phi_{n-1} = -\Delta^2\rho_n \qquad (18)$$

which holds for every internal point n on the mesh. We thus have the set of linear equations:

$$
\begin{aligned}
-2\Phi_1 + \Phi_2 &= -\Delta^2 \rho_1 \\
\Phi_1 - 2\Phi_2 + \Phi_3 &= -\Delta^2 \rho_2 \\
\Phi_2 - 2\Phi_3 + \Phi_4 &= -\Delta^2 \rho_3 \\
&\cdots \\
&\cdots \\
\Phi_{J-2} - 2\Phi_{J-1} + \Phi_J &= -\Delta^2 \rho_{J-1} \\
\Phi_{J-1} - 2\Phi_J &= -\Delta^2 \rho_J
\end{aligned}
\tag{19}
$$

The unknown variables Φ_j for all j may be represented as a vector $\underline{\Phi}$, while the set of equations consists of a known vector $\underline{w}$, and the boundary conditions $\Phi_1 = w_1$ and $\Phi_J = w_J$ specify the value of the potential at the end points; so that:

$$\underline{A}\ \underline{\Phi} = \underline{w} \tag{20}$$

where the matrix $\underline{A}$ has the form

$$
\underline{A} = \begin{bmatrix}
-2 & 1 & & & & \\
1 & -2 & 1 & & \bigcirc & \\
 & 1 & -2 & 1 & & \\
 & & 1 & -2 & 1 & \\
 & \bigcirc & & 1 & -2 & 1 \\
 & & & & 1 & -2
\end{bmatrix}
\tag{21}
$$

A has non-zero elements along only the central three diagonals of the matrix and is termed a <u>tridiagonal</u> matrix. If we choose a difference mesh of great resolution by taking N very large, typically for example N = 10,000, the matrix $\underline{A}$ will consist of many

elements, but it is said to be sparse, in that very few of its elements are nonzero. In addition, it should be noted that, in the particular case of Poisson's equation, the elements along each diagonal have constant values.

Tridiagonal matrices arise in general from considering one-dimensional problems in space. We can formulate now the general boundary value problem in one-dimension:

$$f\frac{d^2u}{dx^2} + g\frac{du}{dx} + hu = w \tag{22}$$

where f(x), g(x), h(x) and w(x) are known functions of x, and we are required to find solutions for the dependant variable u. Agai using difference calculus, on an equally spaced mesh $1 \leq n \leq N$,

$$\frac{f_n}{\Delta^2}(u_{n+1} - 2u_n + u_{n-1}) + \frac{g_n}{2\Delta}(u_{n+1} - u_{n-1}) + h_n u_n = w_n \tag{23}$$

which may be rewritten as:

$$\alpha_n u_{n+1} + \beta_n u_n + \gamma_n u_{n-1} = w_n \tag{24}$$

where
$$\alpha_n = \frac{f_n}{\Delta^2} + \frac{g_n}{2\Delta}$$
$$\beta_n = h_n - 2\frac{f_n}{\Delta^2}$$
$$\gamma_n = \frac{f_n}{\Delta^2} - \frac{g_n}{2\Delta}$$

These equations pertain to a general point in the domain $1 \leq n \leq N$, and boundary conditions are applied at the end points x, and x_N.

In differential form, the general boundary conditions may be written as

$$\left(a\frac{du}{dx}+bu\right)_{x_1}=c \tag{25}$$

$$\left(a'\frac{du}{dx}+b'u\right)_{x_N}=c' \tag{26}$$

where a, b, c, a', b' and c' are prescribed constants. Obviously

(i) if a = 0, the dependant variable u is specified at the end points,

(ii) if b = 0, the derivatives at the end points are defined.

Boundary condition (25) may be written in different form as

$$a\left(\frac{u_2-u_1}{2\Delta}\right)+bu_1=c \text{ or } au_2+(2b\Delta-a)u_1=2c\Delta$$
$$\text{or } \alpha_1 u_2+\beta_1 u_1=\omega_1 \tag{27}$$

or

Similarly boundary condition (26) may be written as:

$$a'\left(\frac{u_N-u_{N-1}}{2\Delta}\right)+bu_N=c'$$
$$\text{or } -a'u_{N-1}+(2b\Delta+a)u_N=2c'\Delta$$
$$\text{or } \gamma_N u_{N-1}+\beta_N u_N=\omega_N \tag{28}$$

where α_1, β_1, ω_1, α_N, β_N and ω_N are defined.

Thus, we have transformed the general differential equation to a set of simultaneous linear equations, with

$$\underline{A}=\begin{bmatrix} \beta_1 & \alpha_1 & & & \\ \gamma_2 & \beta_2 & \alpha_2 & & \\ & \gamma_3 & \beta_3 & \alpha_3 & \\ & & \cdots & \cdots & \cdots \\ & & & \alpha_{J-1} & \beta_{J-1} & \alpha_{J-1} \\ & & & & \gamma_J & \beta_J \end{bmatrix} \tag{29}$$

which is again tridiagonal, but with variable nonzero elements.

If we consider the boundary-value problem in two dimensions in space, the matrix $\underline{A}$ is similarly sparse but is not simply tridiagonal. As an example, let us formulate Poisson's equation in two dimensions and in Cartesian coordinates:

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = -\rho \tag{30}$$

If it is required to solve this equation in, say, the rectangular region R, $0 \leq x \leq X$, $0 \leq y \leq Y$, we may divide the region R by a two-dimensional mesh. We approximate the operator $\frac{\partial^2}{\partial x^2}$ by Δ_x'', $\frac{\partial^2}{\partial y^2}$ by Δ_y''

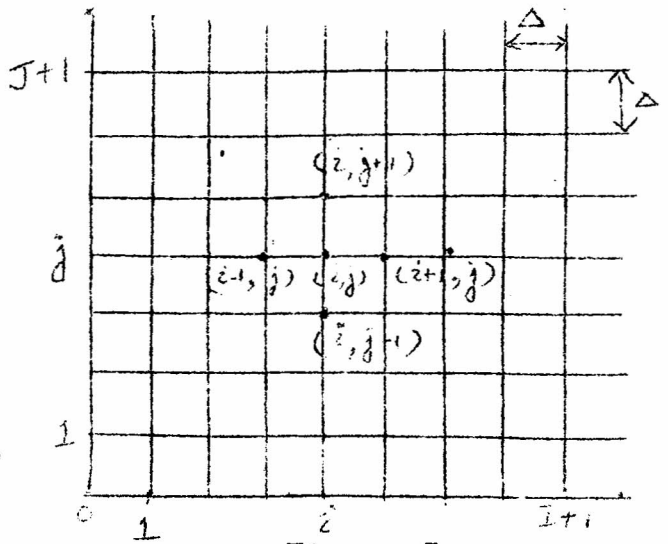


Figure 5

If $\rho_{i,j}$ is defined at every lattice point (i,j) we wish to determine the potential $\Phi_{i,j}$ at every lattice point (i,j) in R, where, from the differential equation, the potential $\Phi_{i,j}$ satisfy,

$$(\Phi_{i+1,j} - 2\Phi_{i,j} + \Phi_{i-1,j}) + (\Phi_{i,j+1} - 2\Phi_{i,j} + \Phi_{i,j-1}) = -\Delta^2 \rho_{i,j}$$

or

$$\Phi_{i+1,j} + \Phi_{i-1,j} + \Phi_{i,j+1} + \Phi_{i,j-1} - 4\Phi_{i,j} = -\Delta^2 \rho_{i,j} \tag{31}$$

at every point (i,j) on the mesh. Let us now replace all the two-dimensional variables by singly subscripted elements:

$$u_k = u_{iJ+j} = \Phi_{i,j}$$
$$\omega_k = \omega_{iJ+j} = -\Delta^2 \varphi_{i,j} \qquad (32)$$

where the vectors $\underline{u}$ and $\underline{\omega}$ have the dimension IJ. The above set of difference equations now read as:

$$u_{k+J} + u_{k-J} + (u_{k+1} - 4u_k + u_{k-1}) = \omega_k$$

The elements of the vector $\underline{u}$ are obtained as solutions to these simultaneous equations and they satisfy the matrix equation:

$$\underline{A}\ \underline{u} = \underline{\omega}$$

where A has the form

$$\left[\begin{array}{cccc|cccc|cccc}
-4 & 1 & & & 1 & & & & & & & \\
1 & -4 & 1 & & & 1 & & & & & & \\
 & 1 & -4 & 1 & & & \ddots & & & & & \\
 & & \cdots & & & & & & & & & \\
 & & & 1 & -4 & 1 & & 1 & & & & \\
\hline
1 & & & 1 & -4 & 1 & & & 1 & & & \\
 & 1 & & & 1 & -4 & 1 & & & 1 & & \\
 & & \ddots & & & \cdots & & & & & & \\
 & & & 1 & & & 1 & -4 & 1 & & 1 & \\
\hline
 & & & & & & & & & & & \\
 & & & & & & & & & & &
\end{array}\right] \qquad (33)$$

Clearly, from this two-dimensional problem, the tridiagonal property of the matrix has been lost, but on the other hand the matrix remains sparse with very few non-zero elements. Matrices arising from problems in 2-dimensional space have the quindiagonal form, though the five diagonals with finite elements are not the leading diagonals. The coupling in two-dimensions is more complex and consequently solutions are more difficult to obtain. As shown in (33) the matrix can be usefully partitioned, where each partitioned matrix relates to different columns of the two-dimensional space mesh. As in the one-dimensional case, in the two-dimensional Poisson equation also, the elements of the matrix on each diagonal are constant. Along these very lines, one can extend the formulation of boundary-value problems to three or more 'space' dimensions but with greater complexities.

We now discuss how the tridiagonal matrix equation (24) can be solved. Let us seek a recurrence solution of the form:

$$u_{n+1} = x_n u_n + y_n \,. \tag{34}$$

where x_n and y_n, in the above algorithm, are intermediate variables to be determined. Substituting (34) for u_{n+1} eq.(24) we have:

$$\alpha_n (x_n u_n + y_n) + \beta_n u_n + \gamma_n u_{n-1} = w_n$$

or

$$u_n = -\frac{\gamma_n}{\alpha_n x_n + \beta_n} u_{n-1} + \frac{w_n - \alpha_n y_n}{\alpha_n x_n + \beta_n} \tag{35}$$

Since the algorithm (34) should also apply at the point n-1,

$$u_n = x_{n-1} u_{n-1} + y_{n-1} \quad , \qquad (34')$$

we have

$$x_{n-1} = -\frac{\gamma_n}{\alpha_n x_n + \beta_n} \quad \text{and} \quad y_{n-1} = \frac{\omega_n - \alpha_n y_n}{\alpha_n x_n + \beta_n} \qquad (36)$$

Equations (34) and (36) provide us a double recursive procedure to solve the matrix equation (24). Notice that the expression on the right-hand side of (34) involves two arithmetic operation (viz. one addition and one multiplication); the expression $\alpha_n x_n + \beta_n$ involves 2 operations, so that x_{n-1} involves four operations (2 operations for the denominator factor, one subtraction and one division) and y_{n-1} involves three operations (one multiplication, one subtraction and one division, the two operations for the denominator factor having been counted earlier). Thus the total number of arithmetic operations required in the expressions for x_{n-1}, y_{n-1} and u_{n+1} are <u>nine</u>.

The boundary conditions at n=N yield the starting values x_{N-1} and y_{N-1} as:

$$x_{N-1} = \frac{\gamma_N}{\beta_N} \quad \text{and} \quad y_{N-1} = \frac{\omega_N}{\beta_N} \qquad \text{(from 28))}$$

Two boundary conditions of frequent occurrence are: (i) the dependant variable u is explicitly defined at the boundary (a' = 0 in eq.26) so that $x_{N-1} = 0$ and $y_{N-1} = u_N$ & (ii) the derivative at the boundary vanishes (b'=0 in eq.26) so that $x_{n-1} = 1$

and $y_{N-1}=0$. With these boundary values, the mesh is scanned completely downwards (n=N to 1) to get all the variables x_n and y_n from (36). These values of x_n and y_n together with the first value of u_1 given by the boundary condition (27) at n = 1, are used in (34) and by scanning the mesh upwards (n=1 to N) we obtain all the values of u_N.

These solutions can be put in matrix form. For this purpose, let us define two operators L_+ and L_- such that:

$$L_+ \{u_n\} = \{u_{n+1}\},$$

and $$L_- \{u_n\} = \{u_{n-1}\} \qquad (37)$$

These ladder operators have the matrix forms:

$$L_+ = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} ; \quad L_- = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & & 1 & 0 \end{bmatrix} \qquad (38)$$

where L_+ is also called as the cyclic matrix and it has the property $L_+^n = I$. Also, since

$$L_+ L_- = I = L_- L_+ , \qquad (39)$$

it follows that L_+ is the inverse of L- and vice versa. In terms of these operators, the matrix equation (24) becomes:

$$\left[P L_+ + Q + R L_-\right] \underline{u} = \underline{w} \qquad (40)$$

where P, Q, R are diagonal matrices:

$$P = \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_{N-1} \\ & & & & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_N \end{bmatrix}, \quad R = \begin{bmatrix} 0 & & & \\ & \gamma_2 & & \\ & & \gamma_3 & \\ & & & \ddots \\ & & & & \gamma_{N-1} \end{bmatrix} \tag{41}$$

As before, we ask for a solution of the type (34):

$$L_+ \underline{u} = X \underline{u} + \underline{y} \tag{42}$$

where X is a diagonal matrix. Multiplying (42) by P from the left and rearranging:

$$[PL_+ - PX + 0]\, \underline{u} = P \underline{y} \tag{43}$$

Subtracting (43) from (40), we have:

$$(Q + PX)\, \underline{u} + R L_- \underline{u} = \underline{w} - P \underline{y}$$

or

$$\underline{u} = -\frac{R}{Q + PX} L_- \underline{u} + \frac{\underline{w} - P \underline{y}}{Q + PX} \tag{44}$$

which is similar to (34'). Eq.(44) is precisely of the form (42) of the solution we asked for, provided:

$$L_- X = -\frac{R}{Q + PX} \quad \text{and} \quad L_- \underline{y} = \frac{\underline{w} - P \underline{y}}{Q + PX} \tag{45}$$

which are matrix forms of equations (36).

We find that the tridiagonal equations can be solved in an extremely simple and efficient manner, without recourse to the straight forward Cramer's rule for matrix inversion, which involves for too many arithmetic operations. Clearly, an n x n tridiagonal matrix equation operating on a vector of dimension n, only $7n$ arithmetic operations are required to solve the system.

The occurrence of tridiagonal matrix equation in a problem in Solid State Physics is discussed by Professor V.Radhakrishnan in his article, in this report. Similar different equations occur in the numerical solution of the Schrodinger equation for a local plus a non-local potential, and the interested reader may refer to the articles of Kermode (1975).

The general matrix equation (2) contains a matrix which is not sparse. In such a case, the following procedure can be employed to obtain an exact solution. Let the system of equations be:

$$\begin{aligned} a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n &= w_1 \\ a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n &= w_2 \\ \vdots \qquad\qquad & \\ a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n &= w_n \end{aligned} \tag{46}$$

Choosing the first of these equations as the <u>pivotal equation</u>, it is straightforward to eliminate u_1 from the remaining (n-1) equations. So that, we have:

$$\begin{aligned} a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n &= w_1 \\ a'_{22}u_2 + \cdots + a'_{2n}u_n &= w'_2 \\ \vdots \qquad\qquad & \\ a'_{n2}u_2 + \cdots + a'_{nn}u_n &= w'_n \end{aligned} \tag{47}$$

where $a'_{ij} = \frac{a_{11}}{a_{i1}} \times a_{ij} - a_{1j}$

and $w'_i = \frac{a_{11}}{a_{i1}} w_i - w_1$ (48)

Choosing the second of the equations in (47) as the pivotal equation, one then proceeds to eliminate u_2 from the (n-2) equations below the pivotal equation. Similar procedures are adopted, using suitable expressions corresponding to (48) until, we obtain the following set of equations:

$$\begin{aligned} a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n &= w_1 \\ a'_{22}u_2 + \cdots + a'_{2n}u_n &= w'_2 \\ a''_{33}u_3 + \cdots + a''_{3n}u_n &= w''_3 \\ \vdots \\ a^{(n-1)}_{nn}u_n &= w^{(n-1)}_n \end{aligned} \quad (49)$$

From this triangular system of equations (49), the variables u_n to u_1 can be successively evaluated, starting with last of the equations which yields u_n straightaway.

In matrix notation, the above Gauss elimination method of solving the matrix equation exactly, involves the factorization of the matrix $\underline{A}$ into two triangular matrices, such that

$$\underline{A} = \underline{L}\,\underline{U} \quad , \quad (50)$$

where the forms of $\underline{L}$ and $\underline{U}$ are:

$$\underline{L} = \begin{bmatrix} \times & & & \\ \times & \times & & 0 \\ \times & \times & \times & \\ \times & \times & \cdots & \times \end{bmatrix} ; \quad \underline{U} = \begin{bmatrix} \times & \times & & \times & \times \\ & \times & \cdots & \times & \times \\ & & & & \\ & & 0 & & \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

These are the lower ($\underline{L}$) and upper ($\underline{U}$) triangular matrices. When these are inverted, they r etain their form. Hence, the solution is by inverting $\underline{L}$ and $\underline{U}$ in turn. So that: $A\,\underline{u} = \underline{w}$ which now reads $\underline{L}\,\underline{U}\,\underline{u} = \underline{w}$ has the solution:

$$\underline{u} = \underline{U}^{-1}\underline{L}^{-1}\underline{w}$$

If the unknown vector $\underline{u}$ is of length n, it can be easily shown that the number of arithmetic operations required to obtain the solution is of the order $\frac{2}{3} n^3$.

Thus, it is important that the choice of method for solving a problem be very carefully made, before resorting to programming or coding the numerical procedure for use on a digital computer.

REFERENCES

1. D.Potter, 'Computational Physics', John Wiley & Sons (1973)
2. L.Fox and D.F.Mayers, 'Computational methods for Scientists and Engineers', Clarendon Press, Oxford (1968).
3. J.C.Taylor, 'Computational Physics', in Phys. Bulletin, 27, 116, March 1976.
4. M.W.Kermode and A.McKerrell, J. Phys. G: Nucl. Phys. 1, 623 (1976); ibid, 1, L40 (1976).
5. V.Radhakrishnan, this Matscience Report.

ON SOME ASPECTS OF INFINITE MATRICES

P.N.Shivakumar
Department of Applied Mathematics
University of Manitoba
Winnipeg, Canada.

****

Acknowledgements:

I wish to thank Professor Alladi Ramakrishnan for inviting me to give a talk and write the review. The review was written during my Sabbatical leave at the Department of Computational and Statistical Science, University of Liverpool. My thanks are also due to Professor L.M.Delves for many discussions and for providing the facilities.

*****

1. INTRODUCTION

1.1 The intention of the present paper is to introduce some of the aspects of infinite matrices and some applications. In mathematical formulation of physical problems and their solutions, infinite matrices arise more naturally than finite matrices. Infinite matrices play an important role in the theory of summability of divergent sequences[2]. The theory of infinite matrices has a colourful history[1] and has its origin in 1884. Infinite determinants were first introduced into analysis in 1886 in the discussion of the Wellknown Hill's equation. In 1906, Hilbert used infinite quadratic forms (which are equivalent to infinite matrices) to solve Fredholm integral equations. Within a few years many of the

theorems fundamental to the theory of abstract operators on function spaces were discovered although they were expressed in special matrix terms. In 1929 John von Neumann showed that an abstract approach was powerful and preferable to using infinite matrices as a tool for the study of operators. The natural development of the theory of function space operators had its origin in the theory of infinite matrices. The applications we will be concerned with here are, linear algebraic systems, linear first order differential systems and conformal mapping. In the above problems, the tools of functional analysis establish in a limited number of cases existence and uniqueness of solutions under stated conditions, but the results are of generally little use in deriving explicit error bounds for approximate solutions or for computational work. Even when one can establish the convergence of the solutions of the truncated systems to the solution of the original infinite systems, explicit error bounds are usually somewhat very difficult to work out. In what follows our main aim is to get meaningful error bounds, in which process existence and uniqueness of solutions fall out as a bonus. It must be stressed however, that the matrices we are concerned with have special stated structures. There is very little work regarding the eigenvalues of infinite matrices [4] except in some very special cases. We give below some examples of special matrices.

1.2 Let $A=(a_{ij})_{i=0\ j=0}^{\infty\ \infty}$ be an infinite matrix for which

(a) $\sum_{k=1}^{\infty} |a_{nk}| \leq M$ for every $n \geq n_0$,

(b) $\lim_{n\to\infty} a_{nk} = \alpha_k$ for every fixed k,

(c) $\sum_{k=1}^{\infty} a_{nk} = b_n \to \alpha$ as $n \to \infty$

Then A is called a Kojima matrix where α_k, α are its characteristic numbers. A has the property that if $\{x_i\}_{i=1}^{\infty}$ is a bounded convergent sequence then $\left\{\sum a_{ij} x_j\right\}_{i=1}^{\infty}$ is also a bounded convergent sequence.

1.3 If $H=\left(\frac{1}{i+j-1}\right)_{i=0\ j=0}^{\infty\ \infty}$, then H is called a Hilbert matrix. For sequence truncation $H^{(n)}$ of H, the inverse is explicitly known[3]. It is also known that the above segments of the Hilbert matrix are highly ill-conditioned with respect to inversion.

We will consider the operator S defined by

$$Sf(s) = \int_0^1 \frac{f(t)}{1-st}\,dt .$$

Considering S as an operator in $L^2[0,1]$, then S has no eigenvalue. In fact Magnus[2] showed that the L^2 spectrum of S is the interval $[0,\pi]$ and it is purely continuous. The operator S is closely related with the infinite Hilbert matrix. For more results concerning the Hilbert matrix please refer to (6). It is known

that 2)

$$f(s) = \frac{1}{\pi} \int_0^1 \frac{f(t)}{1 - st}\, dt \tag{1.3.1}$$

has an exact solution $y(x) = A\sqrt{1-x}\; k\sqrt{x}$ when $\lambda = \frac{1}{\pi}$ and where k is the complete elliptic integral of the first kind. Note that the kernel is not square integrable. An interesting question is whether there are any eigenvalues other than $\frac{1}{\pi}$. Also interesting is the integral equation

$$f(s;a) = \lambda \int_0^a \frac{f(t;a)}{1 - st}\, dt \qquad 0 < a < 1 \tag{1.3.2}$$

and its solution as $a \to 1$. Note that the kernel in (1.3.2) is square integrable. Rewriting (1.3.2.) as

$$f(s;a) = \lambda \int_0^a \left(\sum_{n=0}^{\infty} s^n t^n \right) f(t;a)\, dt$$

we derive by elementary methods

$$y_i^{(a)} = \lambda \sum_{j=1}^{\infty} \frac{a^{i+j-1}}{i+j-1}\, y_i^{(a)}, \quad i = 1, 2, \cdots, \infty \tag{1.3.3}$$

where

$$y_i^{(a)} = \lambda \int_0^a f(t;a)\, t^{i-1}\, dt, \quad j = 1, 2, \cdots \to \infty .$$

Now letting $a \to 1$ formally in (1.3.3) we get

$$y_i^{(1)} = \lambda \sum_{j=1}^{\infty} \frac{1}{i+j-1}\, y_j^{(1)}, \quad i = 1, 2, \cdots, \infty$$

and for which it is well known that $\lambda = \frac{1}{\pi}$ is the spectral radius. It is interesting to note that a sequence of integral equations (1.3.2) with square integrable kernels appear to yield solutions of (1.3.1). A question to answer is whether the systems (1.3.1) and (1.3.4) are equivalent.

REFERENCES

1. BERNKOPF, M.: A history of infinite matrices, Archive for History of Exact Sciences, Vol.4, No.4, 1968, 308.

2. COOKE, R.G.: Infinite Matrices and Sequence Spaces, Dover 1955.

3. DAVIS, P.: Interpolation and Approximation, Blaisdell Publishing Co. N.Y.1965.

4. HANANI, H., NETANYAHU, E. and REICHAW, M.: Eigenvalues of Infinite Matrices, Colloq. Math., Vol.XIX, 1968, 89.

5. MAGNUS, W.: On the Spectrum of Hilbert's Matrix, Amer. J. of Math. 72, 1950, 699.

6. PUTNAM, C.R.: Commutation Properties of Hilbert Space Operators and Related Topics, Springer Verlag, N.Y. (1967).

2. LINEAR ALGEBRAIC SYSTEMS

2.1 In this section we are concerned with existence, uniqueness of solutions of the system

$$(2.1.1)\quad \sum_{j=1}^{\infty} a_{ij}\, x_j = b_i\,, \quad i = 1, 2, \cdots, \infty$$

and also with the error analysis involved in the square truncation of the infinite system given by

$$(2.1.2)\quad \sum_{j=1}^{n} a_{ij}\, x_j^{(n)} = b_i\,, \quad i = 1, 2, \cdots, n\,.$$

In vectorial form (2.1.1) will be written as $A\underset{\sim}{x} = \underset{\sim}{b}$. For a historical account of infinite systems see 3). The following well-known result due to Polya is given in 2).

Theorem: Consider the infinite system of linear equations

$$(2.1.3)\quad \sum_{j=1}^{\infty} a_{ij}\, x_j = b_i\,, \quad i = 1, 2, \cdots$$

where $\{b_i\}$ is an arbitrary sequence and $A = (a_{ij})$ satisfies the conditions:

(C_1) the submatrix formed from the first n rows of (a_{ij}) with the first q columns omitted is of rank n for any n and q;

(C_2) for j = 2,3... we have $\lim_{k\to\infty} \dfrac{a_{j-1\,k}}{a_{jk}} = 0$

Then there exists an infinite sequence $\{x_i\}$ satisfying (2.1.3) with all the lefthand sides absolutely convergent.

Note that no assumptions are made concerning b_i's and the diagonal elements a_{ii}. Condition (C_1) is often difficult to verify. The method of proof precludes the uniqueness of the solution. In fact there must exist infinitely many solutions.

An example of uniqueness of solutions assuming existence is given by Ching and Chiu[1]:

Theorem: Let $A_n = (a_{ij})$, $1 \leq i,j \leq n$ be the nxn matrices obtained from A. Either one of the following conditions is sufficient for the uniqueness of solutions of (2.1.1).

(i) $$\liminf_{n\to\infty} \frac{\prod_{i=1}^{n}\left(\sum_{j=1}^{\infty}|a_{ij}|^2\right)}{|\det A_n|^2} < \infty .$$

(ii) $$\liminf_{n\to\infty} \frac{\left(\sum_{i=1}^{n}\sum_{j=n+1}^{\infty}|a_{ij}|^2\right)\left(\prod_{i=1}^{n}\left[\sum_{j=1}^{n}|a_{ij}|^2\right]\right)}{|\det A_n|^2} < \infty$$

In the above $\{a_{ij}\}_{j=1}^{\infty}$ is assumed to be in l^2 for each i.

2.2 We now discuss a system for which tools of functional analysis can be used to derive existence and uniqueness of solutions and also show that the solution is the limit of the solutions of the system[4]. The method however does not lead to estimates for the solution. By a classical approach, we get existence, uniqueness and meaningful estimates for the solution[5].

Consider the infinite system of linear equations (2.1.1) with $a_{ii} \neq 0$ for all i and

$$\sigma_i |a_{ii}| = \sum_{j \neq i} |a_{ij}|, \quad 0 \leq \sigma_i < 1. \tag{2.2.1}$$

Theorem: Assume

$$\sup_i \frac{b_i}{|a_{ii}|} < \infty \tag{2.2.2}$$

Then the system (2.1.1) with (2.2.1) and (2.2.2) has a unique solution $x = (x_i) \in \ell^\infty$.

Proof: Writing D^{-1} for the diagonal matrix diag $\left[\frac{1}{a_{ii}}\right]$, (2.1.3) implies that $D^{-1} b \in l^\infty$ and clearly $Ax = b$ iff $D^{-1} Ax = D^{-1} b$. Also $I - D^{-1} A$ is a bounded linear operator on ℓ^∞ with $\| I - D^{-1} A \| = \sigma < 1$. Now define a mapping T on l^∞ by

$$Tx = (I - D^{-1} A) x + D^{-1} b \quad (x \in \ell^\infty)$$

Then for two elements $x \in \ell^\infty$, $y \in \ell^\infty$

$$\| Tx - Ty \| = \| (I - D^{-1} A)(x - y) \| \leq \sigma \| x - y \|$$

so that T is a contraction mapping on l^∞. Then there is exactly one element $x \in l^\infty$ such that Tx = x. Since Tx=x is equivalent to $D^{-1}Ax = D^{-1}b$, the Theorem is proved.

As a corollary to the above theorem, we will show that if there are constants $\sigma > 0$ and $M \geq 0$ such that

$$(2.2.3)\quad |a_{ij}| - \sum_{j \neq i} |a_{ij}| \geq \delta, \quad i = 1, 2, \cdots$$

and

$$(2.2.4)\quad \sum_{j \neq i} |a_{ij}| \leq M, \quad i = 1, 2, \cdots$$

then for any $b = (b_i) \in l^\infty$, the infinite system (2.1.1) has a unique solution $x = (x_i) \in l^\infty$.

The corollary will follow from the theorem once we show that (2.2.1) holds. Because of (2.2.3), a_{ii} are bounded away from zero and therefore $b \in l^\infty$ implies $D^{-1}b \in l^\infty$ and also

$$\sum_{j \neq i} \left| \frac{a_{ij}}{a_{ii}} \right| \leq 1 - \frac{\delta}{|a_{ii}|} \qquad \text{for all } i .$$

For those i such that $|a_{ii}| \geq 2M$,

$$\sum_{j \neq i} \frac{|a_{ij}|}{|a_{ii}|} \leq 1 - \frac{\delta}{M}$$

and for those i such that $|a_{ii}| > 2M$, (2.2.4) implies

$$\sum_{j \neq i}^{+} \left| \frac{a_{ij}}{a_{ii}} \right| < \frac{1}{2}.$$

Therefore $\sigma_i \leq \max. \left(\frac{1}{2}, 1 - \frac{\delta}{2M} \right)$ for all i thus implying (2.2.1).

Further it can be shown that the solution to (2.1.1) can be approximated by solutions of finite truncated systems.

In [5] the solution of (2.1.1) under the hypotheses

(H_1) $\sigma_i |a_{ii}| = \sum_{j \neq i}^{+} |a_{ij}|, \; 0 \leq \sigma_i < 1$ (equivalent to (2.1.4)

(H_2) $\sum_{i=1}^{\infty}{}^{+} \frac{1}{|a_{ii}|} < \infty$

(H_3) $\sum_{\substack{j=1 \\ j \neq i}}^{\infty}{}^{+} |a_{ij}| \leq M$ for some M and all i (equivalent to (2.1.5))

(H_4) $\sup_i b_i < \infty$

is discussed.

Note that (H_2) and (H_4) imply (2.1.3). For the truncated system (2.1.2), it is inverted and estimates found. Some of them are

$$(2.2.5) \quad |x_j^{(n)}| \leq \sum_{k=1}^{n}{}^{+} \frac{|b_k|}{|a_{kk}|(1+\sigma_k)} \prod_{k=1}^{n} \frac{1+\sigma_k}{1-\sigma_k}$$

$$(2.2.6)\quad \left| x_j^{(q)} - x_j^{(p)} \right| \leq P \sum_{i=p+1}^{\infty} \sigma_i + Q \sum_{i=p+1}^{\infty} \frac{1}{|a_{ii}|}$$

where P and Q are known constants. From the above relations it is shown that

$$\lim_{n \to \infty} x_j^{(n)} = x_j$$

and that $\{x_j\}$ is in fact a solution of (2.1.1). Under the additional hypothesis

$$(H_5)\quad \sum_{i=1}^{\infty} |a_{ij}| < \infty \qquad \text{for each fixed j,}$$

uniqueness was established by demonstrating

$$AA^{-1} = A^{-1}A = \underline{I}\ .$$

By letting n tend to infinity in (2.2.5) and q tend to infinity in (2,2.6), we have all the needed estimates for the solution of the infinite system.

REFERENCES

1. CHING, CHIN-HUNG and CHUI, CHARLES K. : Uniqueness of solutions of an infinite system of equations, Rocky Mountain Journ of Mathematics Vol.4, No.4, 1974.

2. COOKE, R.G. : Infinite matrices and sequence space, Dover 1955.

3. KANTOROVICH, L.V. and KRYLOV, V.I. : Approximate methods of higher analysis, Interscience, N.Y. 1964.

4. McLURE, J.P. and WONG, R. : Infinite systems of linear equations, (Private communication).

5. SHIVAKUMAR, P.N. and WONG, R. : Linear equations in infinite matrices, Jour. of Lin. Alg. and its Appls., Vol.7, 1, 1973, 53.

3. LINEAR FIRST ORDER DIFFERENTIAL SYSTEMS

3.1. We are interested in the solutions of

$$(3.1.1)\qquad \frac{dx_i}{dt} = \sum_{j=1}^{\infty} a_{ij}\, x_j(t), \quad (x_i(0) = y_i,\ i = 1,2,\cdots\infty$$

where $A = (a_{ij})$ is an infinite matrix. Staying within the framework of abstract spaces, a vector $z(t) = (z_i(t))$ in the finite interval $[0,T]$ will be said to belong to S if $z_i(t)$ is absolutely continuous for every i and

$$\|z(t)\| = \sum_{i=1}^{\infty} |z_i(t)| \leq \gamma < \infty$$

with the sum converging uniformly. We are interested in the solution of (3.1.1) which belong to S.

Bellman[1] established the existence and uniqueness of solutions of (3.1.1) belonging to S when

$$(H_1)\qquad \sum_i |y_i| \leq M < \infty$$

and

$$(3.1.2)\qquad \sum_i \sum_j |a_{ij}| < \infty.$$

Shaw[5] established the existence of solutions of (3.1.1) when (H_1) is satisfied and when

$$(3.1.3)\qquad \sum_{j=1}^{\infty} |a_{ij}| \leq \beta < \infty \quad \text{for all } i$$

and demonstrated uniqueness when

$$(H_2)\quad \sum_{i=1}^{\infty} |a_{ij}| \leq \alpha < \infty \qquad \text{for all } j.$$

The condition (H_2) implies that A is a bounded operator on l_1 since

$$\|A\| = \sup_{\|z\| \neq 0} \frac{\|Az\|}{\|z\|}$$

and

$$\|Az\| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij} z_j|$$

$$= \sum_{j=1}^{\infty} \left\{ \sum_{i=1}^{\infty} |a_{ij}| \right\} |z_j|$$

$$\leq \sup_{j} \sum_{i=1}^{\infty} |a_{ij}| \, \|z\|$$

$$\leq \alpha \|z\|.$$

Considering A as a bounded operator on l_1 it generates a semigroup ex(At) and the solution of (3.1.1) is given by[3]

$$x(t) = \exp(At)y. \tag{3.1.3}$$

By a simple application of Gronwall's inequality, uniqueness of solutions also can be established.

In a recent paper[4] McLure and Wong have proved the existence and uniqueness theorems for the solution of (4.1.1) belonging to S under the assumptions (H_1) and

$$(A_1)\quad \sup\{a_{ij} : i = 1,2,\ldots,\infty\} < \infty$$

and

$$(A_2)\quad \sum_{\substack{i=1 \\ i\neq j}}^{\infty} |a_{ij}| \leq M, \quad j = 1,2,\ldots,\infty .$$

for some finite constant M. They further show that if the matrix is strictly diagonally dominant in the columns

$$|a_{ii}| - \sum_{j\neq i} a_{ji} \geq \delta \quad (j = 1,2,\ldots,\infty)$$

the solution to (3.1.1) is the limit of the solutions to the finite systems obtained by truncating A after n rows and n columns. From their analysis which uses the semigroup theory, a meaningful error analysis for truncated systems does not follow. For a detailed

discussion of the semigroup theory we refer to 3). In the next section we give briefly the error analysis for the solution of the truncated systems of (3.1.1) under abated conditions.

3.2 We are interested in the solutions of (3.1.1) under the conditions (H_1), (H_2) and [2,6]

(H_3) $|a_{ij}| \leq C_i$, for all $j < i$

and

(H_4) $\sum_{i=1}^{\infty} C_i < \infty$.

Proceeding formally, we set

(3.2.1) $$x_i(t) = y_i + \sum_{j=1}^{\infty} b_{ij} \frac{t^j}{j!}$$

for the solution of (3.1.1) and we set

(3.2.2) $$x_i^{(n)}(t) = y_i + \sum_{j=1}^{\infty} b_{ij}^{(n)} \frac{t^j}{j!}$$

for the solution of

(3.2.3) $$x_i^{(n)} = \sum_{j=1}^{n} a_{ij} x_j^{(n)}, \quad x_i(0) = y_i, \quad i = 1, 2, \cdots, n$$

It is easy to see that

$$b_{ij}^{(n)} = \begin{cases} y_i, & j = 0 \\ \sum_{k=1}^{n} a_{ik} b_{k\,j-1}^{(n)}, & j > 0. \end{cases}$$

Using induction, we can establish that for each fixed j and n > m

$$(3.2.4)\quad \sum_{i=1}^{n} |b_{ij}^{(n)}| \leq M\alpha^{j}.$$

$$(3.2.5)\quad \sum_{k=m+1}^{n} |b_{kj}^{(n)}| \leq \alpha^{j} \sum_{k=m+1}^{n} |y_k| + j\alpha^{j-1} M \sum_{k=m+1}^{n} c_k.$$

$$(3.2.6)\quad \sum_{i=1}^{m} |b_{ij}^{(n)} - b_{ij}^{(m)}| \leq j\alpha^{j} \sum_{k=m+1}^{n} |y_k| + \frac{j(j-1)}{2} M\alpha^{j-1} \sum_{k=m+1}^{n} c_k.$$

After some manipulation, we get

$$(3.2.7)\quad |x_i^{(n)}(t) - x_i^{(m)}(t)| \leq \alpha t e^{\alpha t} \sum_{k=m+1}^{\infty} |y_k| + \frac{Mt^2\alpha}{2} e^{\alpha t} \sum_{k=m+1}^{\infty} c_k.$$

By virtue of (H_1) and (H_4), $\{x_i(t)\}_{i=1}^{\infty}$ is a Cauchy sequence, and hence converges to a limit. It can be verified that this limit in fact is the solution to (3.1.1) under the hypotheses (H_1) - (H_4)

Letting $n \to \infty$ in (3.2.7) we get

$$(3.2.8) \quad \left| x_i(t) - x_i^{(m)}(t) \right| \leq \alpha t e^{\alpha t} \epsilon_1 + \frac{M t^2 \alpha}{2} e^{\alpha t} \epsilon_2$$

where for given $\epsilon_1 > 0, \epsilon_2 > 0$, , there exists k such that the above relation holds for $n \geq k$.

3.3. There are several generalizations possible. One can consider the non-homogeneous system

$$\dot{x}_i = \sum_{j=1}^{\infty} a_{ij} x_j(t) + f_i(t), \quad i = 1, 2, \cdots, \infty$$

and also when $a_{ij} = a_{ij}(t)$. It will be interesting to conside unbounded operators on l_p. A problem for consideration would be a set of conditions other than (H_2) that imply A is bounded on l_1 (or l_p). It would be useful to derive the above results when the infinite matrix is diagonally row dominant with the diagonal terms a_{ii} being negative and $|a_{ii}| \to \infty$ as $i \to \infty$.

REFERENCES

1) BELLMAN, R. : The boundedness of solutions of infinite systems of linear differential equations. Duke Math. J. Vol.14, (1947), 695-706.

2) CHEW, K.HL : Finite and infinite matrices and some applications, Ph.D. thesis, University of Manitoba, Canada, 1975.

3) KREIN, S.G. : Linear differential equations in Banach space, Transl. Math. Monographs. Vol.29, A.M.S., Providence, R.I. (1971).

4) McLURE, J.P. and WONG, R. : On infinite systems of differential equation Canadian J. of Math. (Private communication)

5) Shaw, L. : Existence and approximations of solutions to an infinite set of linear time-invariant differential equations. SIAM J. Appl. Math., Vol.22, 2, 1972, 266.

6) SHIVAKUMAR, P.N. and CHEW, K.H. : Infinite matrices in linear differential systems.

4. CONFORMAL MAPPING OF DOUBLY CONNECTED REGIONS

4.1 Solution of a large number of problems in modern technology such as leakage of gas in a graphite brick of gas cooled nuclear reactor[2], analysis os stresses in a solid propellant rocket grain[9], simultaneous flow of oil and gas in concentric pipes[7] hinges critically on the possibility of conformally mapping a doubly connected region onto a circular annulus. If D is a doubly connected region of the z-plane, then the frontier of D consists of two disjoint continua C_0 and C_1. We will assume that each of C_0 and C_1 contains more than one point. Then it is wellknown[1] that D can be mapped one to one conformally onto a circular annulus. Moreover if a and b are the radii of two concentric circles of the annulus, then the modulus D given by b/a is a number uniquely determined by D.

The difficulties involved in finding such a mapping function are described in 3). In fact papers concerning specific regions are few in literature. Hockney[2] considers the region D where the inner boundary is a circle and the outer boundary is a concentric square; while Laura[4] considers the region D with a circular external boundary and an internal boundary which consists of several axes of symmetry. Narodetskii and Sherman[5] discuss the mapping of a region bounded by an ellipse and a circle. Most of the methods use integrals of the Cauchy type and then use truncation procedures to get numerical results. Check

of the numerical results from theoretical considerations is far from satisfactory. See 6), 3) and references in them for numerical work regarding mapping of doubly connected regions. In the next section we will describe a method of reducing the conformal mapping problem to a problem in solution of an infinite system of linear algebraic equations.

4.2 The mapping function[8)]

$$(4.2.1)\quad \omega(z) = \exp\left[\log z + \phi(z)\right],\ z = x + iy = re^{i\theta},$$

which is unique except for an arbitrary rotation maps $D + C_1 + C_0$ onto the annulus $0 < a \leq |\omega| \leq b < \infty$ where the ratio b/a is unique and $\phi(z)$ is regular in D. For convenience, we will assume that the origin in the z plane is not included in the doubly connected region. We will assume

$$(4.2.2)\quad \phi(z) = \sum_{n=-\infty}^{\infty} e_n z^n$$

Hence for all $z \in C_0$, we need

$$(4.2.3)\quad \log(z\bar{z}) + \phi(z) + \overline{\phi(z)} = \log b^2.$$

and for all $z \in C_1$,

$$(4.2.4)\quad \log(z\bar{z}) + \phi(z) + \overline{\phi(z)} = \log a^2.$$

Without any loss of generality, we will assume b to be unity.

If $z = f_1(\zeta)$ maps conformally the simply connected region enclosed by C, to a circle of radius 1 in the ζ plane, then

$$\log[f_1(\zeta)\overline{f_1(\zeta)}] + \phi(f_1(\zeta)) + \overline{\phi(f_1(\zeta))} - 2\log a = 0 \quad \text{on} \quad \zeta\bar{\zeta} = 1 .$$

Using Laurent series expansion for $\phi(z)$, we derive a set of infinite linear equations for the c_n's and a. Similarly if $z = f_1(\zeta')$ maps conformally the simply connected region enclosed by C_0 to a circle of radius 1 in the ζ' -plane, we get another set of infinite linear equations in c_n's. Combining the two sets of equations for c_n's, the existence and the uniqueness of the mapping function depends on the existence and uniqueness of solution of c_n's and a. The above method is illustrated in describing an application in the next section.

4.3 Let

(4.3.1) $$C_0 : x^2 + y^2 = 1 ;$$

(4.3.2) $$C_1 : (x-e)^2 + y^2 = \beta , \quad 0 < e < \beta , \quad e+\beta < 1$$

It is easy to verify that

(4.3.3) $$\phi(z) = c_0 + \sum_{n=1}^{\infty} c_n \left(z^n - \frac{1}{z^n}\right)$$

satisfies (7.2.4) on $z\bar{z} = 1$, provided

(4.3.4) $$c_0 = \log a .$$

Clearly

$$z = c + \beta\zeta$$

maps conformally the simply connected region enclosed by C_0 to a circle of unit radius in the ζ-plane. Hence for all $z \in C_0$,

$$\left\{\log(c+\beta\zeta) + c_0 + \sum_1^\infty c_n (c+\beta\zeta)^{-n} - \sum_1^\infty c_n \frac{1}{(c+\beta\zeta)^n}\right\}$$

$$+ \text{Conjugate} = 0 \text{ on } \zeta\bar{\zeta} = 1 .$$

Using the fact $\left|\frac{c}{\beta\zeta}\right| < 1$, we can rewrite the above as

$$\left\{\log \beta\zeta + c_0 + \sum_1^n \frac{(-1)^{n-1}}{n}\left(\frac{c}{\beta\zeta}\right)^n + \right.$$

$$\sum_{n=1}^\infty c_n c^n \left(\sum_{r=0}^n \binom{n}{r}\left(\frac{\beta\zeta}{c}\right)^r\right)$$

$$\left. - \sum_{n=1}^\infty \frac{c_n}{e^n}\left(\frac{c}{\beta\zeta}\right)^n \left[\sum_{r=0}^\infty \binom{-n}{r}\left(\frac{c}{\beta\zeta}\right)^r\right]\right\}$$

$$+ \text{conjugate} = 0 \text{ on } \zeta\bar{\zeta} = 1 .$$

After some simplification and rearrangement, we get

$$\left\{ \log\beta + c_0 + \sum_{q=1}^{\infty} c_q e^{q} \binom{q}{0} + \sum_{n=1}^{\infty} \left[\frac{(-1)^{n-1}}{n} - \sum_{q=1}^{n} \binom{-q}{n-q} \frac{c_q}{e^{q}} + \sum_{q=n}^{\infty} c_q e^{q} \binom{q}{n} \right] \left(\frac{e}{\beta\xi} \right)^{n} \right\}$$

$$+ \text{conjugate} = 0$$

Equating the powers of $\frac{e}{\beta\xi}$ to zero, we have the infinite system

$$c_0 + \sum_{q=1}^{\infty} c_q e^{q} \binom{q}{0} = -\log\beta,$$

$$\sum_{q=1}^{n} (-1)^{n-q} \binom{n-1}{q-1} \frac{c_q}{e^{q}} - \sum_{q=n}^{\infty} c_q e^{q} \binom{q}{n} = \frac{(-1)^{n-1}}{n}, \quad n = 1, 2, 3, \cdots, \tag{4.3.5}$$

where we have replaced $\binom{-q}{n-q}$ by $(-1)^{n-q}\binom{n-1}{q-1}$. For the problem to be completed, the existence and uniqueness of the solutions to (4.3.5) should be established.

REFERENCES

1) CARATHEODORY, C. : Theory of functions of a complex variable, Vol.2, Chelsea Pub. Co. N.Y. (1956).

2) HOCKNEY, R.W. : A solution of Laplace's equation for a round hole in a square pcg. J. Soc. Ind. Appl. Math., Vol.12, No.1, March 1964, 1.

3) KANTOROVICH, L.V. and KRYLOV, V.I. : Approximate methods of higher analysis, Interscience, N.Y. 1964.

4) LAURA, P.A. : Conformal mapping of a class of doubly connected regions, NASA Tech. Rep. No.8 Catholic University of America, Washington.

5) NARODETSKII, M.Z. and SHERMAN, D.I. : A problem in conformal transformation. PRIKL. Mat. i MEKH, XIV (1950) 209.

6) RICHARDSON, M.K. and WILSON, H.B. : A numerical method for conformal mapping of finite doubly connected regions.

Developments in Theoretical and Applied Mechanics, Vol.3, Pergamon Press, 1967 305.

7) SHIVAKUMAR, P.N. : Viscous flow in pipes whose cross-sections are doubly connected regions. Appl. Sci. Res. 27, April 1973, 355.

8) SYMM, G.T. : Conformal mapping of doubly connected domains.
Numer. Math. XIII, 1969, 448.

9) WILSON, H.B. : A method of Conformal Mapping and the determination of stresses in solid propellant rocket grains.
Rep. No.5-38, Rohm and Haas Co., Alabama, 1963.

OLD INDIAN THEORY OF NUMBERS AND ITS APPLICATIONS IN NUCLEAR AND SOLID STATE PHYSICS

S.D.Sharma
Department of Physics
Panjabi University
PATIALA (INDIA)

***

In olden times Indian mathematicians were successful to develop some important branches of theory of numbers. Theory of numbers had its own independent way of developments in India. We shall primarily discuss here theory of indeterminate equations and some other problems in the field. Contributions of Indian mathematicians of olden times will be compared with the modern developments in these fields. Some applications of very efficient old Indian methods in modern mathematics, nuclear and solid state physics, will also be discussed.

Credit of solving indeterminate equations of first degree ($Ax+B = C$, where A,B,C, are integers) to get integral solutions for the variables x,y goes to Aryabhata of 5th century A.D.. It is to be pointed out that actually for the first time an indeterminate equation of first degree is found solved in Vedangajyotisha (1400 B.C.)[25] where rishi Lagadha finds the nakshatra-ansha at the ending moment of [illegible] taking ansha to be a nonfraction (integer). Also in sulba-sutras[29) 30)] (800 to 500 B.C.) indeterminate equations of first degree (involved in finding area and numbers of bricks to be used in paving layers of special types of fire-places (agni-chitis) were solved, but the theory is not found elaborated by anyone before Aryabhata. In 1624 A.D. Bachet De Mazariac solved first degree indeterminate equation using a

longer and quite inefficient method[6]. In 1673 John Keresy solved y=1,2,3 ----------and wrote down the values of right hand and left hand sides of the equation and wherever any two terms of these series became equal, the minimum integral values of x,y were obtained directly. This process becomes very laborious if the minimum integral values are quite large.

In 1678 Lagrange used Bachet De Mazariac's method. Here A is substracted from B again and again till the least positive reminder is obtained and then the divisor is taken to be dividend and vice versa. The process is continued till remainder = C (= the additive ="Kshepa" in sanskrit) is obtained. Lagrange developed this very method which is similar to the one already developed by Aryabhata. Saunderson (blind from infancy) developed this very method in 18th century A.D. using theory of continued fractions.

Old Indian method is called Kuttaka (pulverizer). In order to understand the formulation it is necessary that the reader may know the valli-algorithm. Consider the equations of the following type

$$S_{r-1} = A_r S_r + S_{r+1} \qquad (I)$$

where r=1,2,3,----n, and given S_{n+1} = a constant number = A_n (say). We let $S_n = 0$ (in order to have minimum solutions of indeterminate equations). Note that $S_{n+1} = S_{n-1}$ and we put $S_{n-1} = A_n$ (for kuttaka problem A_n= Kshepa).

Equation (I) represents n equations. In order to solve these we write $A_1, A_2, A_3, \cdots A_n$ in a column and 0 at the end. This column is called Valli. A_i's are called terms of the Valli. Note that $S_n = 0$ and $S_{n-1} = A_n$. In order to find S_{n-2} we have to multiply A_{n-1} by S_{n-1} and add S_n to it. At every stage the previous result is multiplied by the next higher term in the valli and the lower result in the process valli is to be added to get next s-value. The process will be called Ākunchan (contraction). (This is author's own terminology). The operations are shown here in the annexed valli-vinyasa. If the number of terms in the valli is odd it is called odd valli and if the number of terms is even it is called even valli. Note that the valli algorithm can be used for simultaneous equations of the above type in general. This avoids quite cumbersome eliminations. Here we have put the valli method in a very general fashion. Equations of the type (I) are often encountered in the theory of indeterminate equations (Here A_i's are integers).

Valli terms	ĀKUNCHAN
A_1	
A_2	
A_3	
A_4	
⋮	⋮
A_{n-4}	$A_{n-4}S_{n-4} + S_{n-3}$
A_{n-3}	$A_{n-3}S_{n-3} + S_{n-2} = S_{n-4}$
A_{n-2}	$A_{n-2}S_{n-2} + S_{n-1} = S_{n-3}$
A_{n-1}	$A_{n-1}S_{n-1} + S_n = S_{n-2}$
A_n	$= S_{n-1} = S_{n+1}$
0	$= S_n$

Valli

The indeterminate equation of first degree i.e. the kuttaka equation:-

$$\frac{AX+B}{C}=Y$$

reduces to valli form whose terms are the quotients in finding the H.C.F. of A and C. (As in continued fraction of A/C). The situation can be very easily visualised. (The reader is advised to see ref.4). It may be noted that if the valli is odd the valli will yield x,y. for additive =-B, and if the valli is even the values will be for additive = +B[4].

Aryabhata Bhakaracharya and many other mathematicians solved many interesting problems using valli-method. The equations of the type

$$\frac{AX^n+B}{C}=Y$$

were also solved. The kuttaka formulations were applied in many Astronomical problems. The timings for a given planetary configuration were determined using kuttaka theory and approximations based on this very formulation. (Which were not studied under separate headings but were evidently clear to those mathematicians who used big yuga cycles for such problems) Similar formulations are studied under the heading Diophantine approximations these days. It is of interest to know that the time period of perigee of the sun was determined using kuttaka theory.

Indeterminate equations of Second Degree

In shulba-sutras we find approximations to some surds which seem to be derived from solutions of indeterminate equations of the type $NX^2+1=Y^2$. These texts belong to few centuries B.C. In Greece Diophantus gave rational solutions of about 40 equations of the type[31] $aX^2+bX+c=Y$ K.Proclus discussing Pythagorian triangles gave solutions of some indeterminate equations of second degree. In 7th century A.D. Acharya Brahmagupta gave samās-vyāsa-bhāvanā lemmas for solving the equation $NX^2+A=Y^2$ for integral X,Y. In 11th and 12th century A.D.Jaideva and Bhaskaracharya II solved the equation $NX^2+1=Y^2$ completely using a method called Chakrawal (cyclic) method. Jacobi considers this method to be the best invention in theory of numbers before the time of Lagrange. Dickson praises Brahmagupta very much for his original contributions in this field many centuries before the delayed breakthrough in west in 17th century A.D.[6] C.O.Selenius has praised Chakrawal method very much[21]. Unfortunately after Bhaskaracharya there was almost no development except some applications in various fields. These days the equation $NX^2+1=Y^2$ is called Pell's equation due to mistake by Euler. In fact this was solved by Fermet in 17th A.D. and earlier by Shri Bhaskaracharya in 12th A.D. & by Jaideva in 11th A.D. Hence we will call this equation Jaideva Bhaskara-equation. The equation $NX^2+A=Y^2$ will be called Brahmagupta equation.

In Varga-prakriti formulation we have a Bhavana-lemma which in general can be put as follows:-

If $X_i, Y_i\ (i=1,2,\cdots r)$ are solutions of $NX^2+A=Y^2$ for additives, $A=A_i$ then X_n, Y_n is the solution of the same equation for $A=\prod_{i=1}^{n} A_i$ which can be easily found by comparing rational and irrational terms in the equation:-

$$\prod_{i=1}^{n}(\sqrt{N}X_i \pm Y_i)=\sqrt{N}\,X_n + Y_n$$

If we take the positive sign the operation is called samāsa-bhāvana and if we take negative sign at any stage the operation is called vyāsa-bhāvanā in that step. One can show that for $A_1=A_2\neq 1$, if Y_1/X_1 is convergent of N of order n_1 and Y_2/X_2 is convergent of order n_2 then the two bhāwanās yield convergents of orders (n_1+n_2) and $(n_1\sim n_2)$ respectively. For single operation one can use Vajrābhyāsa (Cross-multiplication) as given by Brahmagupta[5]. This is very handy algorithm for computations of convergents of higher orders (and hence the new solutions). If we call the minimum integer solution X_0, Y_0 for A=1 as the basic set $(X_0, Y_0, 1)$ then one can say that the Vajrabhyasa operation for vajrabhyasa operation by the basic set on (X, Y, A) for additive A yields other solutions of Brahmagupta equation.

Now let us discuss the Cyclic method of Jaideva and Bhaskara II[21] in order to solve the Varga-prakriti equation with A=1.

For $NX^2+A=Y^2$ let $X=1$ and $Y=e$ then the corresponding additive $A=e^2-N$, using samāsabhāvanā with the

other known set (X_1, Y_1, A_1) we get: $X = X_1 e + Y_1$, $Y = NX_1 + Y_1 e$

$$A = A_1 (e^2 - N)$$

Hence we have new set of solution:-

$$X_2 = \frac{X}{A_1} = \frac{X_1 e + Y_1}{A_1} \quad \ldots (I)$$

$$Y_2 = \frac{Y}{A_1} = \frac{N X_1 + Y_1 e}{A_1} \quad \ldots (II)$$

$$A_2 = \frac{A}{A_1^2} = \frac{e^2 - N}{A_1} \quad \ldots (III)$$

Now one can use kuttaka in (I) and get X_2 in parameter k. The parameter k is so selected that $X^2 - N$ is minimum and then we get Y_2 using new additive A_2 the original varga-prakriti equation. This new set is again used to get a solution for a new additive. The process goes on cyclically, till one gets the additive A=1 or 2. Some used kuttaka also for (II) and thus had shortcuts. Bhaskara solved many problems like $67X^2 + 1 = Y^2$.

It can be shown that the number of convergents in Chakrawāl method is very smaller than in Lagrange's and Euler's Methods. C.O. Selenius[21] has tried to correlate ideal continued fractions, Chakrawal Convergents, but it may be cautioned that the minimisations conditions are not the same in both the cases. Thus it can be shown that although the Chakrawal method has deep minimization properties but it is in no way related with ideal continued fractions as claimed by C.C. Selenius.

Valli Method for solving Bhaskara-Equation.

Here we give valli method for quicker operations[26),4)]. We will take a numerical example for developing valli

$$19X^2 + 1 = Y^2$$

In order to be consistent with the previous notations we let

$$x = S_1 \qquad y = S_0 \qquad \therefore 19S_1^2 + 1 = S_0^2$$

so that $S_0 = \sqrt{19S_1^2 + 1}$.

Here the 1st term in the Surd has the nearest root $4S_1$. We can now write

$$S_0 = \sqrt{19S_1^2 + 1} = 4S_1 + S_2 \quad \cdots \quad (4)$$

$$S_1 = \frac{4S_2 + \sqrt{19S_2^2 - 3}}{3} = 2S_2 + S_3 \quad \cdots \quad (5)$$

Here we have taken the square root in the Surd $4S_2$ and added to $4S_2$ and divided by 3 and took only the perfect quotient $2S_2$ and whatever is the remaining part is taken as S_3. Proceeding this way

$$S_2 = \frac{3S_3 + \sqrt{19S_3^2 + 5}}{5} = S_3 + S_4 \qquad (6)$$

$$S_3 = \frac{3S_4 + \sqrt{19S_4^2 - 2}}{2} = 3S_4 + S_5 \qquad (7)$$

$$S_4 = \frac{3S_5 + \sqrt{19S_5^2 + 5}}{5} = S_5 + S_6 \qquad (8)$$

$$S_5 = \frac{2S_6 + \sqrt{19S_6^2 - 3}}{3} = 2S_6 + S_7 \qquad (9)$$

$$S_6 = \frac{4S_7 + \sqrt{19S_7^2 + 1}}{1} = 1$$

Here we take $S_7 = 0$ and then from the above equations we get the Valli:-

$$\text{Valli (even)} \rightarrow \begin{cases} 4 \\ 2 \\ 1 \\ 3 \\ 1 \\ 2 \\ \binom{1}{0} \end{cases}$$

Here tail of the Valli is taken again $\binom{1}{0}$. In fact the tail can be in general of the type $\binom{a}{b}$. This way Valli with higher order solutions in the tail will yield solutions of still higher orders. The Akunchan in order yields: $X = 39, Y = 170$ which is the desired solution. The tail $\binom{1}{0}$ of the Valli yields the least integral solution. It may be noted that for odd Valli the solution is for additive = -1, and for even Valli the solution is for additive = +1. Note also that the incomplete Valli with the tail $\binom{1}{0}$ gives solutions for an additive $= \pm D$ where D occurs in the expression of the form

$$S_r = \frac{A'_{r+1} S_{r+1} \pm \sqrt{N S^2_{r+1} \pm D}}{D}$$

Here D's will be called Hara's. (Divisors). It is clear that if we would have started with $Nx^2 \pm D = Y^2$ there would have been D^2 in the surd and we could stop the udvalana (generation) of the Valli taking $S_{r+1} = 0$ and $S_r = 1$. Thus using tail $\binom{1}{0}$, the partial Valli will yield results for additive $= \pm D$

depending whether the valli is even or odd. Thus in case of N=19 only three terms with tail $\binom{1}{0}$ yield X=3, Y=13 for A=2.

Generating Valli for Jaideva Bhaskara equation (An algorithm)

The valli for N can be generated easily using the following steps:- (1) Suppose the prakriti $N=A_1^2+D_1$, where A_1 is the nearest integer square root of $N(A_1^2<N)$. We write valli terms, R_i's (Sheshas, additional sets of auxiliary numbers), and Haras (Divisors, which will be the respective additives for the partial vallis), in three columns as shown below. First valli term $= A_1$, first remainder (shesha) $= A_1 = (R_1)$ and first Hara $= D_1 = N - A_1^2$.
(2) In order to get next row elements from a previous one, one has to do simple arithmatical operations as follows:-
(a) Let us suppose that we want to get elements of (n+1)th row using three elements of nth row.

(a) $A_{n+1} = \left\{\frac{A_1 + R_n}{D_n}\right\}$ where the symbol $\{-\}$ denotes only the integral quotient taken and

(b) $R_{n+1} = D_n A_{n+1} - R_n$ (c) $D_{n+1} = \frac{N - R_{n+1}^2}{D_n}$

This way one can generate the valli upto hara=1, where it is complete (and actually restarts). One can stop at a stage for respective additive = hara, if so desired. The plan is shown

below:-

Valli-term	R(shesha)	Hara(=additive)
A_1	$R_1 = (-A_1)$	$D_1 = N - R_1^2$
$A_2 = \left\{\frac{A_1 + R_1}{D_1}\right\}$	$R_2 = D_1 A_2 - R_1$	$D_2 = \frac{N - R_1^2}{D_1}$
. . .	. . .	. . .
$A_{n+1} = \left\{\frac{A_1 + R_n}{D_n}\right\}$	$R_{n+1} = D_n A_{n+1} - R_n$	$D_{n+1} = \frac{N - R_n^2}{D_n}$

It can be easily seen that next cycles of full or partial valli if generated and used will yield same results which can be easily obtained using vajrābhyāsa: Also at any stage if one gets hāra=2, one should stop the udvalana of (generating) the valli and use bhāvanā.

It is to be pointed out that comparison of Lagrange's and old Indian methods reveals that Indian methods are faster than those of Lagrange and other European mathematicians. These methods developed many centuries earlier in India.

Extensive methods for 1st and 2nd degree (and even higher degree) indeterminate equations belong to Aryabhata, Brahmagupta and Bhaskracharya Jaideva etc. More difficult problems were solved by Bhaskar and others. These problems will not be discussed here. The equation-

$$aX + bY + c = XY$$

was solved for integral x and y by Bhaskracharya. This is called the Bhāvita equation. (This can also be solved completely by an

intermediatery of Valli algorithm[19]. After Bhaskracharya about five centuries later, this very problem was tackled by Lagrange. Unfortunately the contributions of Indian mathematicians did not get patended under their names, although these formulations were developed by these oriental scholars, many centuries earlier than by the accidental scholars in the west. (See beautiful unbiased remarks by C.C.Selenius Ref.21).

Old Indian mathematicians applied number theoretical formulations in astronomical and other problems using (the concepts of the so called Diophantine) approximations techniques intutively very well clear to them. It is to be pointed out that Valli-Algorithm can be used for computations of Bessel functions Legendre polynomials etc. Here one has to generalise the Valli-algorithm for equations of the type,

$$S_{n+1} = A_n S_n + B_{n-1} S_{n-1}$$

in order to make the algorithm better efficient.

Also it may be pointed out that in power series metho for finding Eigenvalues of Schrödinger wave equations, one can apply Valli-formulation. Termination of Valli terms. (The terms of the tails of Valli) yield the eigenvalues for the problems concerned. For these details author's book in Sanskrit is in the Press[19]. Here we shall discuss some applications of the theory of numbers in nuclear physics and solid state physics. The treatments will promise for better scope of applications.

It may be pointed out that this attempt for applying this branch of pure mathematics is the first of its kind and is very promising too. Such approaches will be very helpful for Physicists like the pure group-theoretic applications in various branches of physics.

If $\nu_i \ (i = 1, 2, \ldots n)$ are the quantum numbers representing a state of a microscopic system. Then an equation of the type

$$f(\nu_1, \nu_2, \nu_3, \ldots \nu_n) = 0$$

obtained on the basis of some physical constraints to the system, will be in general referred to as quantum constraint equation. It is clear that only integer or half integer solutions for $\nu_i \ (i = 1, 2, 3, \ldots n)$ are permissible in such equations. In fact such constraints arise due to the fact that energy is quantised and this results into special phenomena in microscopic dynamics. (for example transitions between specific descret states are possible under a given potential). These constraints are analoge of constraints in a classical system.

As quantum numbers are either integers or half integers so one can easily understand that theory of indeterminate equations may prove a good tool for getting important informations if the equations of Quantum Constraints are tackled taking into consideration this nature of quantum numbers. In fact in finding the degeneracy of a state labelled by many quantum numbers one is encountered with a simple 1st degree indeterminate equation.

In these cases the total number of permissible solutions is the degeneracy of the state. In case of multi-dimensional or multi-particle system one expects more complicated indeterminate equations of 1st degree. For example if a 4-dimensional harmonic oscillator is in state with 3 of its quantum numbers n_1, n_2, n_3 equal (say n) and the 4th remaining being unique equal to n_4, then it is easy to show that the state with $N = \sum n_i = 10 = 3n + n_4$ is 4-fold degenerate. Such problems are similar to the problems of 100 fowls[1] in theory of numbers where one may be interested to get all permissible solutions of the problem. Here one is encountered with the idea of canonical quantum numbers. A simplest situation of this kind is encountered in shell model when Jensen and Meyer changed the labelling defining $\lambda = \frac{N-\ell}{2} + 1$ In fact equations of quantum constraints are mathematical translations of some physical facts like parity, symmetry, invariance of wave functions under rotations etc. Thus an approach on the basis of integral or half integral nature of quantum numbers is all the way justified on physical grounds. Moreover, such an approach is expected to yield overall information about systematics of quantum mechanical systems.

In case of Clebsch Gordon coefficients with one quantum number $= \frac{1}{2}$ the equation $\left| C^{I_1 \; \frac{1}{2} \; I}_{k \; \pm\frac{1}{2} \; (k \pm \frac{1}{2})} \right|^2 = A$ a A rational fraction p/q (say) leads to an indeterminate equation of first degree. Ond can easily show that

$$C^{(qt-1)\,\frac{1}{2}\,(qt-\frac{1}{2})}_{(2p-q)t\,\frac{1}{2}\,(2p-q)t+\frac{1}{2}} = C^{qt\;\frac{1}{2}\;qt-\frac{1}{2}}_{(q-2p)t\;\frac{1}{2}\;(q-2p)t+\frac{1}{2}} = \sqrt{\frac{p}{q}}$$

where t is a canonical quantum number parameter. The equation $\left|C^{I\ 1\ I}_{K_1\ K_2\ K_3}\right|^2 = P/q$ leads to an indeterminate form of 2nd degree in I and K which can be solved using Brahm Gupat's or Bhaskar's theory or the Lagrange's techniques. In case of quandrupole moments under the assumptions of static deformations and symmetric rotating core Wigner Eckart theorem[2] leads to

$$Q(I,K) = C^{I\ 2\ I}_{K\ 0\ K}\ Q_0 \quad \cdots \quad (10)$$

where Q_0 is the intrinsic quadrupole moment of the system. $Q(I,K)$ is the quadrupole moment of a state $|IK\rangle$ which belongs to the rotational sequence with band head K. Under core-rotation the quadrupole moment will be projected to furnish a value with opposite sign if $C^{I\ 2\ I}_{K\ 0\ K} < 0$ the equation $C^{I\ 2\ I}_{K\ 0\ K} = 0$ furnishes a vargaprakriti equation (Brah Gupat's equations or the so called Pell's equation) of the type $NX^2+1=Y^2$ which can be solved for all integral or half-integral I,k. In general for multipole moments one encounters higher degree indeterminate equations in I and K whose solutions with half-integer and integer values only are permissible. This very formulation can be viewed from a different angle also. The indeterminate equations in quantum numbers (which are obtained on physical grounds for quantum mechanical systems), can be thought of as quantum constraints on the quantum mechanical system. The 1st degree equations in general restrict the number of waves comprising the system or sometimes these may represent the contraints resulting

into specific wave patterns for which some specific degrees of freedom are better dominant. For example in case of an axial rotor model of odd-odd nuclei the fact that relabelling of axes leads to no change in the wave function, implies the quantum constraint:

$$K - \Omega_p - \Omega_n = 2N$$

where N=0 or an integer. For symmetric core N=0. Here K, Ω_p, Ω_n are the projections of total nuclear spin I, and those of angular momenta of proton and neutron j_p and j_n respectively. If for a symmetric core nucleus, $\Omega_{p_{max}}$ and $\Omega_{n_{max}}$ are the maximum projections of angular momenta of proton and neutron available in the state basis for the system for some energy range, then the constraint $K = \Omega_p + \Omega_n$ implies that the number of waves becomes constant for all the states with

$$I \geq I_m = K_{p_{max}} + K_{n_{max}}$$

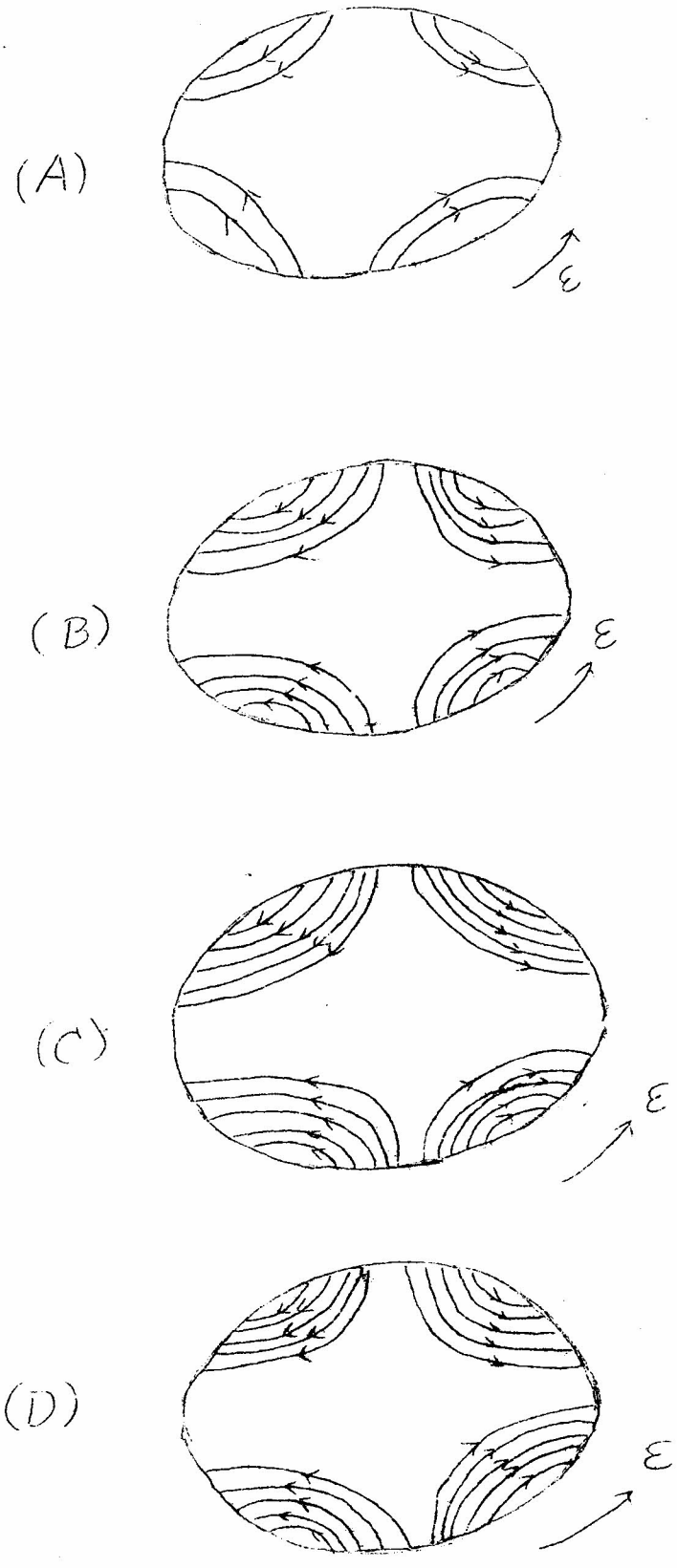


Fig (1)

Figs. depict the increasing number of incoming waves as the nucleus rotates faster and faster. The magnitude of ω shows the extent of spin. The number of waves becomes constant in C & D.

For all I greater than I_m the nuclear core has to furnish one unit of projection of angular momentum in going from I to (I+1) thus the spectrum is better rotational there. For $I < I_m$ nucleons go on contributing more and more which results into more incoming waves on the nuclear surface. Finally the number of waves becomes constant and the system becomes dynamically better rotational. At this stage although the components of state vectors do change on going to higher and higher spins but their dimensionality remains constant. These situations are depicted in figures A, B, C and D. One can think of classical analogue example of rotating sphere having two springs attached to it. When the springs get streched to their maximum the system gets better rotational. It is to be emphasised that if in the above example of a rotating odd-odd nucleus, the particles get excited and thus start furnishing more states to the system, the rotational features start getting masked. These two phenomena of wave saturation and state gradient could explain the features of perturbed rotational bands, systematics of static moments, transition probabilities and the failure of coupling rules in case of odd-odd deformed nuclei throughout the periodic table[15]. Also these concepts could clarify the existence of a new good quantum number in collective approach[18]. As the first degree quantum constraint equations could explain these systematics in nuclear dynamics it is quite believable that higher degree indeterminate equations in quantum numbers will furnish important informations about nuclear dynamics like the vanishing of some properties (nuclear moments etc.) and selection rules for some transitions as will be clear

from our formulation here. It is to be noted that the quantum constraints being discussed here are some physical conditions imposed on the system under the assumptions of the model. In fact these constraints are representing the competition between the non-rotational and rotational tendencies of the system, which are represented by the total spins I (and j in case of particles) and their projections (K or Ω) furnished along symmetry axes, note that the ordinary selection rules too can be thought of as simple quantum constraints which are representing the law of conservation of angular momentum and the availability of the same to the outgoing gamma rays in the processes of transitions. The triangularity check-up conditions for vanishing of Clebsch Gordon coefficients etc. are based on this very fact. One more point is to be emphasised that partly constraints expressed in vanishing of a Clebsch-Gordon coefficient as

$$C^{l_1\ l_2\ l_3}_{0\ \ 0\ \ 0} = 0 \quad \text{if } l_1 + l_2 + l_3 \text{ is odd}$$

Here we will encounter more interesting quantum constraints.

In order to solve eq.(1) for all integral or half-integral. I and K the bhāvanā theorem[3),4),5),6)]. The computations are facilitated using the Vajrabhyas techniques. Here one gets the states $|0\ 0\rangle, |1/2, 1/2\rangle, |3, 2\rangle, |15/2\rangle, |25/2\rangle$ and we stop the bhavana operations at this stage as higher spin states are not as yet expected in nuclear spectra. (The next higher state with vanishing quadrupole moment is found to be K=28 I=48).

Thus in these bands below and above the respective rotational member states, quadrupole moment has opposite sign. In general one can show that Q changes sign at states lying between $I=(3K-1)/2$ and $(4K-1)/2$ (taken to be nearest integers or just half integers as the case may be). Also in case of a single particle

$$Q = -\frac{3K^2 - I(I+1)}{\sqrt{I(I+1)(2I+1)(2I+3)}} \langle r^2 \rangle$$

Thus one can get all the half integer quantum number states for which the quadrupole moment vanishes. A case of many nucleons may be of much interest (Ref. Sach's text appendix[20]).

Using the recursion relation between $C^{I\,K\,I+1}_{K\,0\,K}$ and $C^{I\,K\,I+1}_{K\,0\,K}$ one can easily get an expression for $C^{I\;3I}_{K\,0\,K}$. Thus the octupole moment operator can be proved to be proportional to

$$T_{30}(I) = 3(7/4\pi)^{1/2} I_0 (5I_0^2 - 3I^2 + 1)$$

whose eigenvalues equated to zero yield the equation

$$K[3I(I+1) - (5K^2+1)] = 0$$

Here $I=K=0, \frac{1}{2}, 1$ and using $15K'^2 + 21 = Y^2$ where $K'=2K$ Vyasbhavana yields $K=5$ $I=6$ for even A nuclei and using Vyasbhavan between $(\frac{1}{2}, 6, 21)$ and $(4, 31, 1)$ for $60K^2+1 = Y^2$ one gets $K = 27/2$ $I = 21/2$ for odd-A nuclei and infinite number of high spin states at present of no interest to nuclear physicists. It is evident that in other bands the sign of octupole moment will

change at about the state with Spin I = (3K-1)/2 (Taken to be nearest integer or just half integer as the case may be).

Using the recursion relation one can show that $C^{I\,4\,I}_{K\,0\,K} = 0$ yields

$$105K^4 - 7\{3I^2(I+1)^2 - I(I+1)\} - 5\{6I(I+1) - 5\}\times \{3K^2 - I(I+1)\} = 0$$

which can be solved by reducing it to the Vargaprakriti form. Defining $R = I(I+1)$ one gets

$$K^2 = \frac{5(6R-5) + \sqrt{480R^2 - 660R + 605}}{70}$$

The quantity under redical sign has to be perfect square. $(=Y^2$(say) and letting Y =10t one gets $R = \frac{33+n}{46}$, where

$$460t^2 - 1911 = n^2 \quad \text{---------} \quad (11)$$

which has abhavya roots $(3, 41/23)$ and 7/2. The auxiliary Bhaskar equation for Samas-Vyasbhavana operation on (11) is

$$460T^2 + 1 = m^2$$

which can be solved using Chakrawal method of Valli method or Lagrange's method of continued fractions. Here Valli is

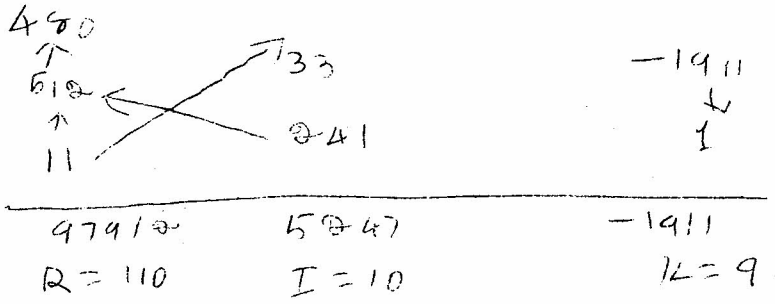
(दर्शनी)

$$\begin{cases} 2 & 1 \\ & 1 \\ & 9 \\ & 1 \\ & 1 \\ & 0 \end{cases}$$

$241 = m$

$11 = T$

It is clear that for t=5/2, m=33 (which can be easily visualised using the solutions corresponding to I=K). Here for I=K=0, Y=25, $t=2\frac{1}{2}$, n=33. Hence using Vyasbhavana in order to be confined to lower spin states one has



One can use other abhvya-roots also as basic sets for bhavana. All these operations prove that no half-integer solutions exist for this equation. Higher spin states are of no interest to nuclear physicists. Thus it is proved that for odd-A nuclei the hexa-decapole moment is never zero except at some bandheads, but in case of even-even or odd-odd nuclei, it is zero at I=10 in the K=29 band. It can be shown that it changes sign at I of the order of 3K in other bands.

It is to be pointed out that higher multipole moments can be studied on similar grounds but it is evident that the equations $C^{I\,L\,K}_{K\,0\,K} = 0$ with $L > 4$ have very rare permissible roots. Thus higher multipole moments will be rarely zero in states other than the bandheads.

It is to be emphasized that these Clabsch Gordon coefficients vanish at these states in spite of the fact that these confirm to the tringuality conditions, but the off-diagonal ($I_i \neq I_f$) Clebsches do not vanish and thus the respective gamma ray

transition probabilities do not vanish even if the respective nuclear moment projects to a zero value in the initial and final states of transitions. For transition probabilities to vanish, the respective intrinsic moment has to vanish. The off-diagonal Clebsch Gordon coefficient will be zero only through parity considerations (although triangularity conditions may hold, for example $C^{I \to I+1}_{0\,0\,0} = 0$). Thus we conclude that in rotational bands isolated states (the states wherefrom or to which the transitions are forbidden) are not possible, but in case of vibrating rotors such states may exist (as in transitions probabilities for states in such nuclei, coefficients like $C^{I<I}_{\nu\,0\,\nu}$ may occur)[7)8)]. The E_2 transitions like $|3_2\rangle \to |2_2\rangle$ in vibrational or asymmetric core nuclei are forbidden and these states are E2 inactive in these transitions. Such states may exist also in beta decay to odd-odd nuclei where transitions probabilities are proportional to squares of certain racha coefficients. Zeroes of some racha coefficients are being evaluated (for example W(2222, 03) = 0 etc. are evaluated) and the presence of such states is under investigation which may be following these special selection rules.

If one considers effective moments of inertia due to any type of mixing caused by cariolis coupling (for $\Delta K=1$ states) or asymmetry effects (for $\Delta K=2$ states) etc. one arrives at expressions which involve the Brahm Gupta expression in the denominator[9]. It is clear that the effective value of moment of inertia will overflow for I=K=0,1 and for K=6, I=8 states etc. but never for odd-A nuclei.

Near these states due to increase in rotational energy deformation is expected to increase which may cause enhancement in E2-transition possibilities. Sometimes there may be termination of rotational bands also.

At present data on nuclear moments for higher members of rotational bands are not reported. Thus the above results at present cannot be verified directly. Bands upto K=15/2 or 9 and spins upto I=25/2, 10 are either observed or expected to be observed in case of high Z even A or odd-A nuclei[10]. The band K=9 with I=10 state is observed in case of H_f^{180} but results on hexadecapole moments of these members of rotational sequence are still awaited[11),12),13)] but it is clear that here we have arrived at better theorems on vanishing of electric, quadrupole and magnetic octupole moment etc. Also we have now the overall picture of variation of these moments in rotational sequences. The projection behavious is well understood this way (variation of magnitudes and signs of projected multipole moments etc. ard now known to a better extent).

It is clear that in the states I=K=0, 1 & $K=6, I=0$ the effective moments of inertia blow up. In the states I=K=0, there is odd-even shift due to coriolis coupling and many such cases are available in odd-odd nuclear spectra. The E2-transitions nearby these states are to be investigated. In case of Pr^{144} the E2-transition $2 \rightarrow 0$ is found enhanced by a factor of the order of 100 in comparison with shell model estimates[14]. In case of Lu^{174}, N_p^{238} and B_k^{250} bands with K=6 are observed but I=8 state is missing[15]. In case of $B_{K=6, I=9}$ state is

observed but there is no state with spin I=8. In these regions enhancement of E2-transitions is expected which will be verified whenever experimental reports are available.

It is to be pointed out that whenever a nuclear moment vanishes the respective spin-spin interaction energy (like quadrupole interaction energy) also vanishes as in the relevant expression for energy we have[2)]

$$\langle I_i I_j \mid \mathcal{H}(I,J) \mid I_i I_j \rangle \sim C^{I\,I\,I}_{I_i\,0\,I_i}\, C^{J\,I\,J}_{I_j\,0\,I_j}$$

Thus the systematics of variations in 2^l-pole and 2^0 pole interaction energy can be inferred easily from such a study.

Similar approach in other fields also can furnish important informations. For example in an admixture of waves one can easily understand the relative prominent contributions of some states using the equation of the type $\left| C^{I_1 I_2 I_3}_{I_1 I_2 I_3} \right|^2 =$ A rational fraction. This equation is an indeterminate equation of 1st and second degree for $I_2 = \frac{1}{2}$ and 1 respectively and has infinite number of permissible solutions but for higher I_2-values it may have infinite or a large number of roots in some cases only. This mathematical fact indicates some special characteristics of dipole terms etc. in interaction energy.

Also such an approach furnishes important information about the Schrodinger waves for these states. In reoriented procession theory for a quadrupole-quadrupole interaction[16)] we have time-dependence of amplitude $e^{i\omega_E (3I_z^2 - I^2 - I)\, t}$ where ω_E

is the frequency of a different type of precession caused by quadrupole moment. This shows that the amplitude becomes constant for the Vargaprakriti states already discussed.

Canonical quantum numbers occur sometimes in effective physical quantities (like effective moments of inertia where one can define an overall rotational quantum number etc.)., or sometimes as indeterminate quantum numbers in a coupling of many angular momenta. The concept can be used for labelling of states and the behaviour of good quantum numbers can be understood in a better way.

In j^n (or l^n) configuration of n fermions or bosons one can study the states using the expression $\nu = J(J+1) - n j(j+1)$ also in atomic theory of hydrogen atom using recursion relations for expectation values of various powers of radii of electronic orbits[17] we get the relation:-

$$\langle r^2 \rangle = \frac{1}{2}\{5n^2 + 1 - 3l(l+1)\} n^2 a^2$$

which is an indeterminate expression in n and l. It is evident that integral roots of the equation $5n^2 + 1 - 3l(l+1) = 0$ will give $l > n$ which is contradicting our hypothesis. Thus $\langle r^2 \rangle$ is always positive. Its variations from state to state can be understood in a better way using such approach.

It is realised that formulations using Markov spectral studies[22] and also Minkowski geometrical studies of numbers[23] will lead to more informations in nuclear dynamics and other microscopic systems.

In lattice structure studies related with lattice coordinates, the problems can be haldled using Minkowsky geometry of numbers[23]. Various theorems on M-body problems etc. have already been developed by Minkowsky, Haris Hancock and others. Studies of conics with integer descrete points will prove very helpful, Lattice-defects can be studied using such a formulation. Necessary backgrounds are already prepared in the detailed developments done by Minkowski. For example maximum number of faces of a maximal M-body (a body or polyhedron figure, which has each face containing at least one lattice point) is residue number in the congruences of the type[23]

$$p \equiv p_0 (\mathrm{mod}\ 2),\ q \equiv q_0 (\mathrm{mod}\ 2),\ r \equiv r_0 (\mathrm{mod}\ 2)$$

This yields 8 residues as can be easily visualised using kuttaka. Thus the maximum number of faces of a maximal M-body will be 2(8-1)=14. (as origin is excluded from these points.) Such a fourteen-faced figure is obtained if all the six vertices of a tetrahedron are sawn off. Also the maximum number of lattice points on the boundary of an M-body can be shown to be the number of solutions of indeterminate equations of the type[23]

$$p \equiv p_0 (\mathrm{mod}\ 3),\ q \equiv q_0 (\mathrm{mod}\ 3),\ r \equiv r_0 (\mathrm{mod}\ 3)$$

The integer nature of the variables restricts the number to $3^3 = 27$ and excluding origin we get the maximum number of lattice points to be 26. Also very general complicated theorems already developed by Minkowski for n-dimensional geometry can be utilised for futher developments. Work in this direction is in progress.

From all these examples we understand that such studies based on number theoretical approaches to indeterminate equations can prove very useful, and theory of quantum constraints can be developed. These studies can be very much helpful for better understanding of many phenomena in various branches of microscopic dynamics.

REFERENCES

1. Needham Joseph "Science and Civilization of China" Mathematics and the Sciences of the Heavens and the Earth. Vol.III. Published by Cambridge University Press (London) 1959.
2. M.E.Rose "Theory of angular momentum" published by John-Wiley & Sons, New York.
3. S.D.Sharma, "Old Indian theory of indeterminate equations" (to be published)
4. Sh.Bhaskracharya "Beej ganitam" (in Sanskrit) a part of "Siddhanta-Shiromani" available from Moti Lal Banarsi Dass, Jawahar Nagar, Bangalow Road, Delhi-7. India.
5. Sh.Brahamgupta "Brahmgupta Siddhanta" (in Sanskrit) available from Moti Lal Banarsi Dass, Jawahar Nagar, Delhi-7 India.
6. Dickson "History of Theory of Numbers", Chesley Publishing Co., New York.
7. J.P.Davidson & M.G.Davidson N,P,33 664 (1962).
8. C.E.Avelodo & J.P.Davidson P.R.No.5, Vol.2, 1934 (Nov.1970)
9. J.D.Singh & S.D.Sharma, Matscience N.P.Symp. report 1973.
10. 'Tables of isotopes' Lawrence Radiation Lab. Univ. of California Berkeley, U.S.A.
11. M.A.Preston "Physics of the Nucleus" Addison Wesley Pub. Co. Inc. New York.

REFERENCES

1. Needham Joseph 'Science and Civilization of China', Mathematics and the Sciences of the Heavens and the Earth. Vol.III Published by Cambridge University Press (London) 1959.
2. M.E.Rose 'Theory of angular momentum' Published by John-Wiley & Sons, New York.
3. S.D.Sharma, 'Old Indian theory of indeterminate equations' (to be published).
4. Sh.Bhaskracharya, 'Beej ganitam' (in Sanskrit) a part of 'Siddhanta-Shiromani' available from Moti Lal Banarasi Dass, Jawahar Nagar, Delhi-7.
5. Sh. Brahmgupta 'Brahmgupta Siddhanta' (in Sanskrit) available from Moti Lal Banarsi Dass Jawahar Nagar, Delhi-7.
6. Dickson 'History of Theory of Numbers', Chesley Publishing Co. New York,
7. J.P.Davidson & M.G.Davidson, N.P.33, 664 (1962).
8. C.E.Aveledo & J.P.Davidson, P.R.No.5, Vol.2, 1934 (Nov.1970)
9. J.D.Singh & S.D.Sharma, Matscience, N.P.Symp. Report 1973.
10. 'Tables of Isotopes, Lawrence Radiation Lab., Univ. of Calif. Berkeley.
11. M.A.Preston, 'Physics of the Nucleus' Addison Wesley Publ. Co. Inc. New York.
12. 'Quadrupole moments of excited states', Physics Abstracts, 73-39062.
13. D.Cline, 'Static magnetic and electric moments of excited nuclear states' University of Rochester Report UR-NSRL-37, 170.
14. S.D.Sharma, BARC, Nuclear Physics Report, 1972
15. S.D.Sharma, Ph.D. Thesis on 'Collective Models of odd-odd nuclei.' Univ. of Kansas, Lawrence (KS)U.S.A.
16. Measurements of Qua.. Moments using reorientation precession Technique' Uppsala University, (Sweden) Publication 1974.

17. Albert Messiah; 'Quantum Mechanics' Published by North Holland Co. Amsterdam.

18. S.D.Sharma, 'A new good quantum number in collective model' BARC Symp. Report, 1973.

19. S.D.Sharma, Bhautiki Ganitam, (in Sanskrit) (in Press) with financial aid from Ministry of Education.

20. R.C.Sachs, Nuclear Theory, Addison Wesley Pub. Co. Cambridge,

21. C.O.Selenius, On Rationale of Chakrawal of Jaideva and Bhaskar II' (Uppsala Univ.). Historia Mathematica, Vol.2/May 1975 (Toronto Univ. Publication).

22. R.T.Bamby, On Markov Spectrum in 'Diophentine Approximations and its applications', Edited by Charles F.Osgood, Academic Press, New York and London 1973.

23. Harris Hancock, Development of Minkowski Geometry of Numbers' Vol.1,2, Dover, New York Publications Inc. New York

24. ShAryabhata 'Aryabhatiyam' Edited by Baldeva Mishra, Published by Bihar Research Society Patna, India.

25. Rishi Lagadha, Vedanga Jyotisham, Commentary by Shri Sham Shastri, Published by Mysore Ass, Supdt. Govt. Branch Press 1936.

26. Sh.Sudhakara Dwivedi Appendix in 'Karana Prakasha', Available from Moti Lal Banarasi Dass, Jawahar Nagar, Delhi.

27. Niven and Zuckerman, 'An Introduction to theory of Numbers' Published by John Willy & Sons, Inc. New York & London.

28. Journal fur Mathematics 3, 1828354, Werke 1-21.

29. Katyayana Shulbsutram, Edited by Vidyadhara Sharma, Kashi.

30. Baudhayana Shulbasutram, Edited by S.Prakasha and R.S.Sharma, Delhi. (1968)

31. Neughebauer Otto (Brown University Rhodes Islands) Exact Sciences is Antiquity (1962).

ON POINT PATHOS GRAPHS OF A GRAPH

(Graphs defined on (0,1) matrix)

Bhalachandra Gudagudi
MATSCIENCE, The Institute of Mathematical Sciences
Madras-20. (INDIA)

***

INTRODUCTION:

Let $A = [a_{ij}]$ be a (0,1) matrix. The graph G(A) of the matrix A is defined as follows. The point set of G(A) consists of all a_{ij}'s with $a_{ij}=1$. Two such points are adjacent if and only if they appear either in the same row or in the same column of A.

The idea of the matrix graphs was provided by the well known theorem of Konig-Egervary[2] on (0,1) matrices.

THEOREM: (Konig-Egervary). Let A be a (0,1) matrix of size m x n. The minimum number of lines that cover all the 1's in A is equal to the maximum number of 1's in A with no two of the 1's on a line.

In ((3) p.2) Hedetniemi indicated that the matrix graph thus constructed has its point independence number equal to the maximum number of 1's in the matrix with no two of the 1's on a line.

In [1] C.R.Cook has defined the clique-vertex graph as one of the special case of the matrix graphs and discussed some of its properties.

The concept of a pathos of a graph G was introduced by F.Harary as a collection of minimum number of line disjoint open paths whose union is G . The path number $P_n(G)$ of a graph G is the number of paths in a pathos. R.G.Stasston, D.D.Cowan and L.O.James have calculated the path number for certain classes of

graphs like trees, cubic graphs complete graphs etc. L.Lovasz[6] has also produced some good bounds for P_n G .

The point-pathos matrix A of G is a pxk matrix $[a_{ij}]$ where $a_{ij} = 1$ if the point v_i belongs to the path p_j in P where P is the set of pathos in G and $a_{ij} = o$ otherwise.

Let P be a pathos of a graph G. The point pathos graph $_pP(G)$ of G is the matrix graph of point-pathos matrix of G. That is the point set of $_pP(G)$ is the set of ordered pairs (p_i, v_j) where v_j is a point of G on the path p_i in P and two ordered pairs (p_i, v_j) and (p_m, v_n) are adjacent if and only if either $p_i = p_m$ or $v_j = v_n$.

Recall that in general path number of any graph is not determined. However, path number is determined for tree and some special classes of graphs.

In this paper we mainly deal with point pathos graph $_pP(T)$ of a tree T and obtain its characterisation. We also show that the row graph of the point pathos matrix of a graph G is isomorphic to the pathos graph P(G)[7]. And the clique graph (or block graph) of a column graph of the point pathos matrix is also isomorphic to the pathos graph P(G).

Let T be a tree and P be a set of pathos of T. The rows of the point pathos matrix of T correspond to the paths in P of T and the columns correspond to the points of T. The column sums of this matrix correspond to the number of pathos in P, in which a point appears.

In figure 1 a tree T and the set of paths in P are given. In table 1, the point pathos matrix corresponding to the tree T and paths in it are given. In figure 2, the point pathos graph P(T) of a tree (T in figure 1) is given.

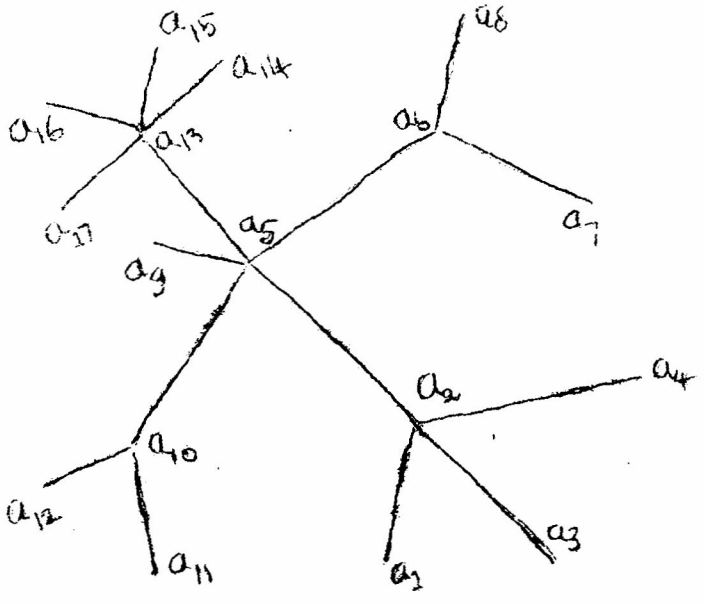


Figure 1

P_1 : a_1 a_2 a_3

P_2 : a_4 a_2 a_5 a_6 a_7

P_3 : a_6 a_8

P_4 : a_5 a_{13} a_{14}

P_5 : a_{12} a_{10} a_5 a_9

P_6 : a_{10} a_{11}

P_7 : a_{16} a_{13} a_{15}

P_8 : a_{17} a_{13}

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}
p_1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p_2	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
p_3	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
p_4	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
p_5	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0
p_6	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
p_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
p_8	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1

Table 1

Point pathos matrix of T in figure 1

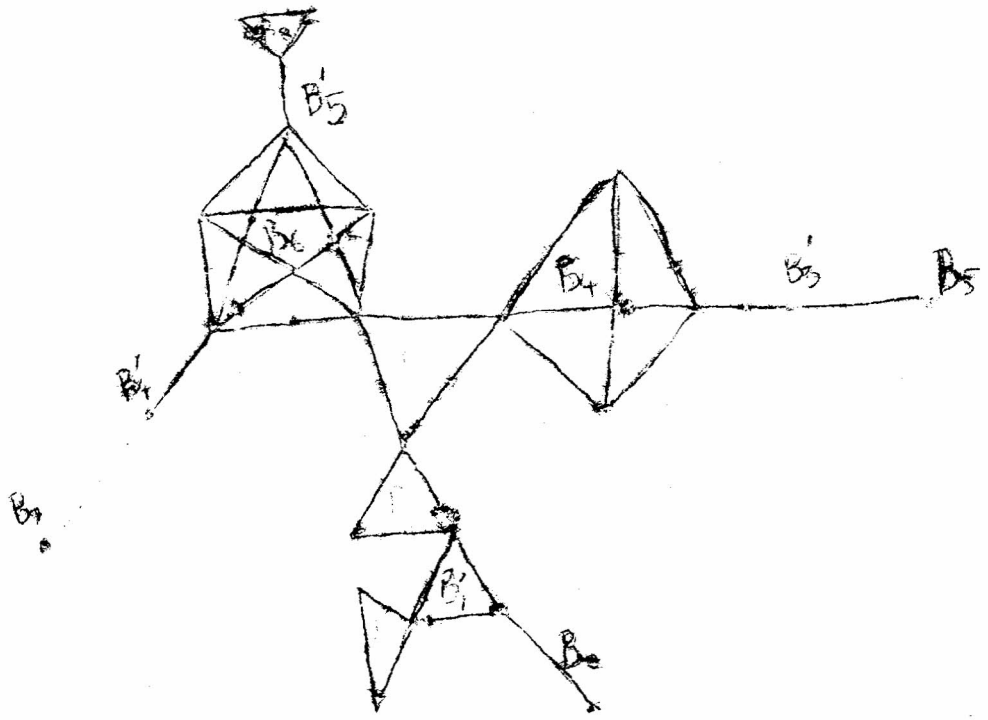


Figure 2

Point pathos graph $_pP(T)$ of a tree T in figure 1.

THEOREM:1 Let T be a tree. Then the number of points in $_pP(T)$ is $\sum_i n_i$, where n_i is the path number of a point v_i in T.

PROOF: If n_i is the path number of a point v_i in T then v_i lies on n_i paths in P. Thus the total number of 1's appearing in the i-th column of the point pathos matrix of G is n_i. And hence total number of 1's appearing in that matrix is $\sum n_i$. This is the number of points in $_pP(G)$. This proves the theorem.

THEOREM:2. Let T be a tree with $pn(v_i)$ as the path number of a non pendant point v_i in T and n_i as the length of a path p_i of P in T. Then the total number of lines in $_pP(T)$ is given by

$$\sum_i n_i \frac{(n_i+1)}{2} + \sum_i \frac{pn(v_i)(pn(v_i)-1)}{2}$$

PROOF: In a point pathos matrix of a tree, the number of 1's is every row corresponds to the number of points in the corresponding path p_j of P in T. And by definition of point pathos graph, points corresponding to 1's in a row induce a complete block in ${}_pP(T)$. If n_i for i = 1,2,.....,n are the lengths of the paths p_i in P of T, then the number of lines in the corresponding complete blocks in ${}_pP(T)$ will be clearly

$$\sum_i \frac{(n_i+1)\ (n_i+1)-1)}{2}$$

$$\text{i.e.} \qquad \frac{n_i(n_i+1)}{2} \qquad \text{(A)}$$

Similarly, the points corresponding to the number of 1's in each column form a complete block in ${}_pP(T)$, which corresponds to the number of paths p_i in P having v_i as a common point. This is clearly path number $pn(v_i)$ of the point v_i. So again the number of lines corresponding to such complete blocks will be

$$\sum_i \frac{pn(v_i)\ \left(pn(v_i)-1\right)}{2} \qquad \text{.....} \qquad \text{(B)}$$

From (A) and (B) it follows that the total number of lines in ${}_pP(T)$ is

$$\sum_i \frac{n_i(n_i+1)}{2} + \sum_i \frac{pn(v_i)\ \left(pn(v_i)-1\right)}{2}$$

This proves the theorem.

Let $A = \{a_{ij}\}$ be a (o,1) point pathos matrix with m rows. Then the points of the row graph

R(A), $\{u_1, u_2, \ldots\ldots, u_m\}$, correspond to the rows of A and two points of R(A) are adjacent if and only if there is a column of A with 1's in the two rows corresponding to the points.

THEOREM: 3. The row graph R(A) of a point pathos matrix of a tree T is isomorphic to P(T), the pathos graph of the tree T.

PROOF: Clearly, there is a one-one correspondence between the points of the row graph of the point pathos matrix and the points of the pathos graph P(T) of the three T. Two points in the row graph of the point-pathos matrix of the tree T, are adjacent if and only if the two rows have a 1 in the same column. This is equivalent to saying that the corresponding paths of the pathos have a point in common. This implies that the row graph R(A) is the interaction graph of the pathos P of the tree T. This proves that the row graph of the point pathos matrix of the graph T is isomorphic to its pathos graph P(T).

Example. The following graph in figure 3 is the row graph of a tree T shown in figure 1.

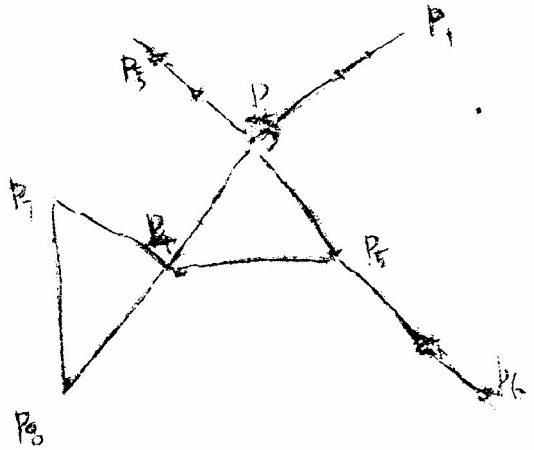


Figure 3

Let $A = [a_{ij}]$ be a (o,1) point pathos matrix with n columns. Then the points of the column graph C(A), $\{v_i, v_j, \ldots, v_n\}$ correspond to the columns of A and two points of C(A) are adjacent if and only if there is a row of A with 1's in two columns corresponding to the points.

THEOREM:4 Let T be a tree and P be the set of pathos in T. Then the clique graph (or block graph) of the column graph of the point-pathos matrix is again isomorphic to the pathos graph P(T) of the tree T.

PROOF: By the definition of the pathos graph P(G) of a graph G, there is one-one correspondence between the paths in P of G and the points in P(G).

By the definition of a column graph there is a one-one correspondence between the complete blocks of C(A), the column graph of point-pathos matrix A, and the paths p_i in pathos P of tree T.

Also, by the definition of the clique graph, there exists one-one correspondence between the cliques (or blocks) of G and the points of clique graph (or blocks) of G and the points of clique graph (or block graph).

Now, since every block in the column graph C(A) of a tree T is complete, it follows that there exists one-one correspondence between the paths p_i in P of T, and hence the number of points in the corresponding pathos graph P(T) of a tree T and the points of the clique graph (or block graph) of the column graph of the point pathos matrix of the tree T.

On the other hand any two points are adjacent in P(T) implies that the corresponding paths p_i, p_j in P of T have a point in common. Consequently this implies that the complete blocks (or cliques) Q_i, Q_j in C(A) corresponding to p_i, p_j in P of T, have a point in common. This in turn implies that the two points in the clique graph (or block graph) of a column graph of a point pathos matrix, corresponding to Q_i, Q_j are adjacent. Thus it follows that clique graph (or block graph) of the column graph C(A) of the point pathos matrix of a tree T is isomorphic to the pathos graph P(T). This proves the theorem.

Example.

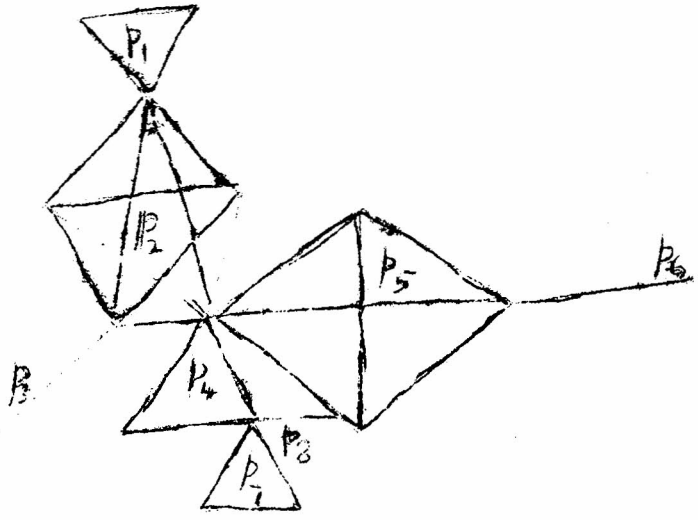


Figure 4

The graph in the above example (see figure 4) is the column graph C(A) of a point pathos matrix of a tree T in figure 1. It is easy to observe that clique graph (or block graph) of this graph C(A) is isomorphic to the pathos graph P(T) of the tree T.

The following Corollary follows from the above theorems.

Corollary 4.1: Let T be a tree and A be its point pathos matrix. Then the row graph R(A) of A is isomorphic to the clique graph (block graph) of its column graph C(A).

CHARACTERIZATION OF $_pP(T)$ OF A TREE T

The following two theorems characterize the point pathos graphs of a tree T.

THEOREM 5: A graph G is point pathos graph $_pP(T)$ of some path T if and only if G is complete.

PROOF: NECESSITY. Any path T of length n contains (n+1) points and has pn(T) = 1. So the point pathos matrix contains one and only one row and has as many number of columns as there are points in T. Naturally this matrix has all 1's in the same row. This implies clearly that $pP(T)$ is complete on (n+1) points.

SUFFICIENCY: Suppose G be a complete graph, say K_n. Then it is easy to see that the graph obtained by removing all chords and a line of K_n, is a path T of length n-1, which has a path number unity and n points. Clearly K_n is $_pP(T)$ of T thus obtained. This proves the theorem.

THEOREM 6: A graph G is the point pathos graph of a tree T(Which is not a path), i.e. $G = {}_pP(T)$ if and only if

(a) Every block of G is complete.

(b) There exist two subsets Π_1 and Π_2 of the set Π of block of G, such that

i) $\pi = \pi_1 \cup \pi_2$

ii) No two blocks belonging to the same set have a point in common.

iii) At each point of any block belonging to one of the subsets say π_2, there is a block belonging to π_1.

PROOF: NECESSITY. Let G be a point pathos graph of a tree T. Consider the point pathos matrix A of T. By the definition of G, the 1's in A correspond to the points of A and every block in G is complete. Also, each column (or row) in A corresponds to a complete block in G, and conversely. The order of a complete block in G is equal to the number of 1's appearing in the corresponding column (or row) of A.

Let A be a mxp matrix and

$$\pi_1 = \{B_1, B_2, B_3, \ldots\ldots, B_m\}$$

be the set of blocks of G corresponding to the rows of A. Also let $\pi_2 = \{B'_1, B'_2, B'_3, \ldots\ldots, B'_p\}$ be the set of blocks of G corresponding to the columns of G. It is easy to see that no two blocks in π_1 or in π_2 have a common point.

Now, consider B'_1 of the set π_2, and let its order be r. This block corresponds to the first column in A and there are "r" 1's in the first column. Let

$$i_1, i_2, i_3, \ldots\ldots, i_r$$

be the rows of A_1 which have 1's common with the first column of A. Then each of the blocks

$$B_{i_1}, B_{i_2}, B_{i_3}, \ldots\ldots, B_{i_r}$$

of G have a common point with B'_1. Likewise we can show that every block in $\mathfrak{B}$ satisfies the condition (ii) of (b). This proves that the conditions in the theorem are necessary.

SUFFICIENCY: Let G be a connected graph satisfying the given conditions in the theorem..

Let $\mathfrak{B}$ be the set of blocks of G and $\mathfrak{B}_1, \mathfrak{B}_2$ be two subsets of $\mathfrak{B}$ satisfying the given conditions of the theorem. Thus at each point for every block B' in there is a block belonging to $\mathfrak{B}_1$.

Let G' be the graph obtained from G by contracting every block belonging to $\mathfrak{B}_1$ into a single point. Clearly G' is also a graph in which every block is complete. Hence, by an earlier theorem (Theorem 6, Chapter I) G' is a pathos graph of some tree T. Now the construction of the tree T can be undertaken in two ways.

Case 1. If the tree T does not contain points of degree 2, then T can be easily constructed from G' by using the method mentioned in theorem 6 of Chapter I.

Case 2. Suppose the tree contains points of degree 2. In this case we can construct the required tree T' as follows.

Let G" be the graph obtained from G by contracting the blocks belonging to $\mathfrak{B}_2$ into a single point. Then replace every complete block of order n($n \geq 3$) by a cycle of length n. (This can be

easily done by removing all chords of every complete block of order n ($n \geq 3$). Then further remove one line from every cycle of order n ($n \geq 4$) which is not incident to any star-point.

In case the complete block is a triangle, then remove the line which is adjacent to an even star-point.

Also, in a case in which each of the point of the complete block is incident to a block, then remove the line that is incident to two points of even degree suitably.

Let T' be the tree thus obtained. Now one can easily verify that the point pathos graph ${}_pP(T')$ is the graph G. This proves the sufficiency and hence the theorem.

Now, by the help of the following example we can show that the condition (b) of the theorem is irredundant.

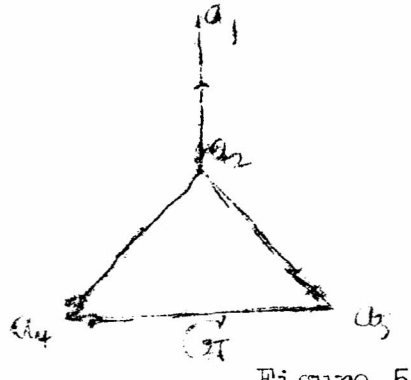


Figure 5

Here we observe that every block in G is complete but it is not a point pathos graph of any tree.

REFERENCES

1. Cook, C.R. — Graphs associated with (0,1) arrays, The University of Iowa Themis Project, Technical Report 28 (1970).

2. Hall, M., Jr. — Combinatorial Theory, Blaisdell, Waltham, Mass., (1967).

3. Hedetniemi, S.T. — Graphs of (0,1)-matrices, The University of Iowa, Themis Project Technical Report 14 (1969).

4. Harary, F., — Covering and Packing in Graphs - I Annals of the New York Academy of Sciences, 175 (1970) pp 198-205.

5. Stanton, R.G., Cown, D.D. James, L.O., — Proceedings of the Louisiana Conference on Combinatorics, Graph Theory and Computing (1970). pp 112-135.

6. Lovasz, L., — On covering of graphs Theory of graphs Proceedings of the colloquium, held at TIHANY, Hungary, 1966, edited by P.Erdos and G.Katona-

7. Gudagudi, B.R., — "Some topics in Graph Theory", Ph.D. Thesis, Karnatak University, Dharwar.

SAVING COMPUTER TIME IN SOLVING SYSTEMS OF LINEAR EQUATIONS BY ITERATIVE METHODS

M.N.Channabasappa
Department of Mathematics,
Karnataka Regional Engineering College,
Surathkal. (India)

*****

1. INTRODUCTION

The importance of the problem of solving systems of linear equations in scientific investigations can hardly be over-emphasised. This is very well pin-pointed by Dahlquist and Djörock [1] when they remark that the solution of a linear system of equations enters in at some stage in about 75 per cent of all scientific problems. Two methods that are widely used to compute solutions of linear systems are: (i) the Gauss elimination method, which is a direct method and (ii) the Jacobi's method or its improved version namely the Gauss-Seidel method, which is an iterative method. Iterative methods are preferred to direct methods in situations where the systems to be solved are very large and are with sparse coefficient matrices.

The purpose of this paper is to show how computer time can be save in solving linear systems of equations by iterative methods by using iterated iteration, a technique developed in this paper. An iterative algorithm, when used on a computer, terminates either after it has gone through the prescribed number of times or when the quantity under investigation converges with prescribed accuracy. In either case it is necessary that the computer programme involving

an iterative algorithm makes use of control statements like the DO or IF. For a given iterative process, we will define, what we will call, the iterated iteration formulae of different orders, and will show that the use of such formulae will result in the reduction of the number of times that certain necessary control statements have to be gone through, and consequently, in the saving of computer time.

2. ITERATED ITERATION FORMULAE

Let $f : x_n \rightarrow x_{n+1}$ be a single point convergent iteration function. We define the iterated iteration function of order k associated with f by $g : x_n \rightarrow x_{n+1}$, where $g = f^k = f^h f^{k-h}$, h and k being positive integers, $h \leq k$. We assume that f itself is not obtainable through a similar process, and we therefore, call it a basic iteration function or the iterated iteration function of order one. The iteration formula involving the iterated iteration function of order k will be called the iterated iteration formula of order k or simply, the formula of order k.

We apply the above to Jacobi's method of solving systems of linear equations.

Let the linear system to be solved consist of say, n equations in n unknowns which can always be put in the form

$$\underline{X} = A\underline{X} + \underline{F} \qquad (1)$$

We assume that $\rho(A) < 1$ (2)

where $\rho(A)$ is the spectral radius of A.

Jacobi's algorithm for the system (1) is

$$X^{(n+1)} = A X^{(n)} + F, \qquad n = 0, 1, 2, \cdots \tag{3}$$

The condition (2) is necessary and sufficient[2] for the convergence of Jacobi's algorithm for any arbitrary $X^{(0)}$.

The iterated iteration algorithm of order k for Jacobi's method is

$$X^{(n+1)} = B X^{(n)} + G \tag{4}$$

where

$$B = A^k, \quad G = (A^{k-1} + A^{k-2} + \cdots A + I)F \tag{5}$$

In view of (2), it follows that

$$\rho(B) < 1, \tag{6}$$

which ensures the convergence of the k^{th} order algorithm for any choice of $X^{(0)}$.

3. CHOICE OF THE ORDER k

The discussion in this section is quite general and applies to any iteration process.

Let N_k denote the number of iterations required by a given iterative algorithm to converge with some prescribed accuracy when the order of the formula employed is k, and let M_k denote some measure of the arithmetic involved in one such iteration.

Since the derivation of the formula of order k involves the repeated use of the first order formula at least k times, we can write

$$M_k \geq k\, M_1 \tag{7}$$

$$N_1 \leq k\, N_k \,, \tag{8}$$

the equality in either case holding good when the expression for x_{n+1} in the basic iteration formula contains x_n only once. From (7) and (8) we get

$$M_k N_k \geq M_1 N_1 \tag{9}$$

Clearly, the number of iterations is less for a higher order formula than for the basic iteration formula. Consequently, the number of times that certain necessary control statements have to be gone through is also less for a higher order formula than for the basic iteration formula. This results in the saving of computer time.

In view of (9), the computer time spent on the arithmetic when a higher order formula is used is greater than or equal to the corresponding time when the basic iteration formula is used.

But in view of the fact that the execution time of control statements is much more than the execution time of arithmetic statements (when they are simple enough), the above two opposing factors can combine to still result in the saving of computer time. Thus the technique of saving computer time consists in choosing the iteration formula of the suitable order. This does

not, of course, mean that we can make the time saved arbitrarily large by choosing k arbitrarily large for the reason that when the accuracy is prescribed and when we choose k arbitrarily large, it may be that we will be doing more work than actually necessary for attaining the specified accuracy, and hence the time taken may exceed the necessary minimum.

Hence, for a prescribed accuracy, there will exist an optimum k for which the time saved will be a maximum. This optimum k, for any given algorithm and for any prescribed accuracy has to be determined experimentally.

4. A NUMERICAL EXAMPLE

We give below the results of computation for the determination of optimum k for Jacobi's algorithm applied to the following problem of computing the solution of

$$x_1 = 0.3\,x_1 - 0.2\,x_2 + 0.1\,x_3 - 0.5$$

$$x_2 = 0.5\,x_1 + 0.1\,x_2 + 0.1\,x_3 + 1.7$$

$$x_3 = 0.4\,x_1 - 0.2\,x_2 + 0.2\,x_3 - 0.8$$

with the starting approximation $x^{(0)} = (0,0,0)$.

The computations were carried out on the computer ICL 1909 installed at the Indian Institute of Technology, New Delhi. In order to get perceptible time difference, the processes of different orders were repeated 50 times. The control statement used was the IF statement, and was used, after each iteration, to check whether the prescribed accuracy was attained. Table 1 gives the results of computation.

Table 1 - Results of Computation

Accuracy prescribed	Solution obtained	No. of times the process was re-peated.	k	N_k	Time taken in seconds	Optimum k	percentage of time saved
5 D	x_1=-1.20183		1	14	29		
	x_2=1.08256	50	2	7	22	2	24.1
	x_3=-1.24770		3	5	24		

The results of table 1 confirm the existence of iterated iteration formula of optimum order. The computer time saved by using the optimum order formula is quite substantial (24.1%).

It is to be conceded that programs for optimum order formulae will be iteration formulae. But with bulk memories available in modern computers, this is no serious constraint at all.

The iterated iteration can certainly be used for saving computer time in other iteration algorithms. The results of computation pertaining to (i) the fixed point algorithm applied the solution of Kepler's equation and (ii) the Newton's algorithm for computing square roots are given in the appendix.

REFERENCES

1) Dahlquist, G., Bjorck, A., Numerical Methods, Prentice Hall, 1974, pp.137.

2) Ibid, pp.197.

APPENDIX

Table 2 given below contains the results of computation pertaining to the following two problems:

Problem 1: To find, using fixed point iteration, the root of

$$x = 0.8 + 0.2 \sin x$$

starting with $x_o = 0.8$

Basic iteration formula: $x_{n+1} = 0.8 + 0.2 \sin x_n$.

Problem 2: To compute, using Newton's method, the square root of 41 starting with the approximation $x_o = 6$.

Basic iteration formula : $x_{n+1} = x_n + \frac{r_n}{2 x_n}$

where $r_n = 41 - x_n^2$.

Table 2 - Results pertaining to Fixed Point Iteration and Newton's Algorithm

Problem	Accuracy	Solution obtained	No. of times the process was repeated	k	N_k	Time taken in seconds	Optimum k	Percentage of time saved
1	10 D	0.9643328877	200	1	12	13	3	30.7
				2	7	10		
				3	5	9		
				4	4	9		
2	10 D	6.4031242374	1000	1	4	21	2	38.0
				2	2	13		
				3	2	15		
				4	2	19		

NUMERICAL COMPUTATION OF SURFACE AND COULOMB ENTRIES OF AN AXIALLY DEFORMED NUCLEUS*

V. Devanathan
Department of Nuclear Physics
University of Madras
Madras-600025. INDIA.

****

Introduction

In problems of nuclear fission and heavy ion collisions, one needs to compute the surface and coloumb energies of an axially deformed nucleus. The asymmetric two-centre molecular shape appears to be the most realistic nuclear fission and it can be described by a set of five parameters [1]. In such cases, one has to resort to numerical methods to compute the surface and coloumb energies. The most convenient procedure, as pointed out by Lawrence [2], is to use the Gaussian integration formula [3] This has a precision of order 2n - 1 when only n evaluation points are used.

The surface energy

It is comparatively simply to calculate the surface area of a deformed axially symmetric nucleus. In the case of fission process, it is reasonable to assume conservation throughout the intermediate stage but some compression may take place during the heavy ion collisions.

The surface energy of an axially symmetric nucleus is given by

$$E_S = 2\pi \int \rho \, ds \qquad (1)$$

where

$$ds = dz \sqrt{1 + (d\rho/dz)^2} \qquad (2)$$

To describe the shape fully, ρ should be known as a function of Z and the integration dz should be carried out throughout the length of the nucleus.

The Coulomb energy.

To compute the coulomb energy, first divide the nucleus into thin discs. The next step is to calculate the interaction energy between any two discs A and B as shown in Fig.1. The coulomb

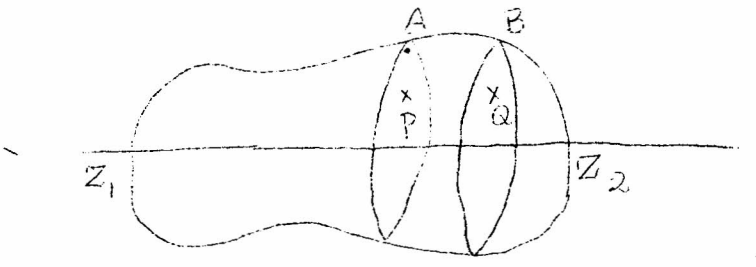


Fig.1

energy is the sum of interaction energies between all such discs and care should be taken to avoid double counting.

Consider a small element of volume $r_b\, dr_b\, d\phi_b\, dz_b$ at Q located on the disc B and find out the potential dv_p at any point P on the disc A.

$$dV_P = \frac{\sigma r_b\, dr_b\, d\phi_b\, dz_b}{d} \qquad (3)$$

where σ is the uniform charge density and d the distance between the two points P and Q

$$d^2 = r_a^2 + r_b^2 - 2 r_a r_b \cos(\phi_b - \phi_a) + (z_b - z_a)^2$$

The potential at P due to the entire disc B is obtained by integration

$$V_p = \sigma\, dz_b \int_0^{\rho_b} \int_0^{2\pi} \frac{r_b\, dr_b\, d\phi_b}{d} \qquad (4)$$

The integration energy between the two discs A and B can be calculated by first computing the coulomb energy due to a small element of charge located at P and then integrating throughout the volume of the disc A.

$$E_{AB} = \sigma\, dz_a \int_0^{\rho_a} \int_0^{2\pi} V_p\, r_a\, dr_a\, d\phi_a \qquad (5)$$

The coulomb energy of the nucleus is obtained by summing all such interaction energies without double counting. This done by integrating dz_b over the entire length of the nucleus Z_1 to Z_2 but restricting the integration of dz_a from Z_1 to Z_b.

There are in total six integrations but ultimately they can be reduced to three integrations. The occurence of d in the denominator causes some difficulty. But it is overcome by the use of formulae involving Bessel functions. We list below the formula used in the reduction.

(i) $$\frac{1}{(R^2+Z^2)^{1/2}} = \int_0^{\infty} e^{-|z|\lambda} J_0(\lambda R)\, d\lambda \qquad (6)$$

(ii) $$J_0(\lambda R) = J_0(\lambda r_a)\, J_0(\lambda r_b) + 2 \sum_{n=1}^{\infty} J_n(\lambda r_a)\, J_n(\lambda r_b) \cos n\phi \qquad (7)$$

where $R = (r_a^2 + r_b^2 - 2 r_a r_b \cos\phi)^{1/2}$

(iii) $$\int_0^{\rho} r\, J_0(\lambda r)\, dr = \frac{\rho}{\lambda} J_1(\lambda \rho) \tag{8}$$

This can be proved by using the recursion relation

$$z J_1'(z) + J_1(z) = z J_0(z) .$$

The Watson's identity

(iv) $$\int_0^{\infty} e^{-|z|\lambda} J_1(\lambda \rho_a) J_1(\lambda \rho_b) \frac{d\lambda}{\lambda^2} \tag{9}$$
$$= \frac{\rho_a \rho_b}{2\pi z} \int_0^{\pi} {}_2F_1\left(\tfrac{1}{2}, 1, 2; -\frac{\omega^2}{z^2}\right) \sin^2\phi\, d\phi$$

where $\omega^2 = \rho_a^2 + \rho_b^2 - 2 \rho_a \rho_b \cos\phi$

(v) $${}_2F_1\left(\tfrac{1}{2}, 1, 2; -\frac{\omega^2}{z^2}\right) = \frac{2|z|}{|z| + \sqrt{z^2+\omega^2}} \tag{10}$$

Using formula (i), we replace $1/d$ by an integral involving the Bessel function and by means of formulae (ii) and (iii), we perform the integration over $d\phi_b$ and dr_b to obtain

$$V_p = 2\pi\sigma\, dz_b \int_0^{\infty} e^{-|z_b - z_a|\lambda} J_0(\lambda r_a) \times \tag{11}$$
$$\times \frac{\rho_b}{\lambda} J_1(\lambda \rho_b)\, d\lambda$$

The integration over $d\phi_a$ and dr_a can also be performed without difficulty and we obtain

$$E_{AB} = 4\pi^2\sigma^2 dz_a dz_b \rho_a \rho_b \int_0^\infty e^{-|z_b - z_a|\lambda} \times \times J_1(\lambda\rho_a) J_1(\lambda\rho_b) \frac{d\lambda}{\lambda^2} \tag{12}$$

Using the Watson's identity, we get

$$E_{AB} = 4\pi\sigma^2 \rho_a^2 \rho_b^2 dz_a dz_b \int_0^\pi \frac{\sin^2\phi \, d\phi}{|z_b - z_a| + \left(|z_b - z_a|^2 + \omega^2\right)^{1/2}} \tag{13}$$

Let $\phi = \pi\omega$

Then

$$E_{AB} = 4\pi^2\sigma^2 \rho_a^2 \rho_b^2 dz_a dz_b \int_0^1 \frac{\sin^2\pi\omega \, d\omega}{|z_b - z_a| + \left(|z_b - z_a|^2 + \omega^2\right)^{1/2}} \tag{14}$$

where

$$\omega^2 = \rho_a^2 + \rho_b^2 - 2\rho_a\rho_b \cos\pi\omega$$

Summing over the interaction energies due to all the discs, we get the coulomb energy Ec of the nucleus

$$E_c = 4\pi^2\sigma^2 \int_{z_1}^{z_b} \rho_a^2 dz_a \int_{z_1}^{z_2} \rho_b^2 dz_b \times \times \int_0^1 \frac{\sin^2\pi\omega \, d\omega}{|z_b - z_a| + \sqrt{|z_b - z_a|^2 + \omega^2}} \tag{15}$$

We are now left with only three numerical integrations and they can be performed using the Gaussian integration formula.

The Gaussian integration formula

The Gaussian integration formula is stated as follows:

$$\int_a^b F(x)\, g(x)\, dx = \sum_{r=1}^{n} \omega_r\, g(x_r) + R_n \qquad (16)$$

In the above ω_r's are the weight functions and x_r's are the associated points. These quantities are independent of the function g(x) but they are dependent on F(x). The reminder R_n is zero if g(r) is a polynomial of order less than 2n. Thus the n-point quadrature formula is exact for all polynomials g(x) whose degree is less than 2n-1.

The limits a and b can be changed into -1 and +1 by the following transformation

$$x \rightarrow \frac{b+a}{2} + \frac{b-a}{2}\, x \qquad (17)$$

Further, if we put F(x) = g(x) = 1, then we find that the sum of weight functions should add up to 2.

$$\sum_{r}^{n} \omega_r = 2$$

For F(x) = 1, it is shown that the x_r's are the zeros of the Legendre polynomials of order n. The corresponding weight functions are positive and can be obtained in terms of P_1's. When n is odd, one of x_r's is zero. If x_r is a root of $P_n(x)$,

then $-x_\gamma$ is also a root. The roots and the weight functions are available to an accuracy of 15 decimal places for values of n upto 16 [4].

In the evaluation of coulomb energy, one has to use only the even quadrature since at $x_\gamma = 0$, one of the factors in the integrand becomes zero/zero and hence cannot be evaluated in a computer. To avoid this difficulty, we use only even quadratures which do not contain the root $x_\gamma = 0$.

Acknowledgements.

Useful discussions with Dr.P.R.Subramanian and Mr.G.Shanmugam are acknowledged with thanks.

References

1. J.Maruhn and W.Greiner, Z. Phys. 251 (1972) 431
2. N.P.Lawrence, Phys. Rev. 139B/1227 (1965)
3. A.D.Booth, Numerical methods, Butterworthe, London (1957)
4. Lowan, Davids and Levenson Bull. Amar. Math. Soc. 48 (1942) 739.

POSITRON WAVE FUNCTION IN SOLIDS*

K. Iyakutti and V. Devanathan
Department of Nuclear Physics, University of Madras,
Madras-25. (INDIA)

****

1. Introduction

In order to obtain a theoretical estimate of the positron annihilation in solids one needs to calculate the positron as well as electron wave functions. Inside the solid the conditions differ widely from those in free space and hence one has to take into account all these facts when calculating the wave function. Standard methods such as APW, KKR, OPW etc are available for the treatment of electrons in solids. Now let us discuss the case of positrons in solids. We note that approximately 98% of the positrons reach the end of their 'path' without suffering annihilation. At the end of their path, the positron K.E. is of the same order of magnitude as that of the fastest electrons in the degenerate electron gas in the metal. Subsequent energy losses for the positron are then possible largely through inelastic collisions with the thermal lattice vibrations of the metal atom. We see that as a result of such inelastic collisions, the positron attain thermal equilibrium with the lattice in a time short compared with the mean time for annihilation. Once thermalized, the positrons diffuse randomly through the metal without any further energy gain or loss on the average and are eventually annihilated. To treat the positrons, the following methods are generally used: 1) The use of analytical function with an

* Presented by Mr.K. Iyakutti.

adjustable parameter in it. 2) Representation of the positron by a δ function. 3) Wigner-Sietz approximation method 4) Plane wave expansion method 5) Pseudopotential calculation. For solids of high atomic number (Z) the problem may become still more complicated. For electrons it is established that the relativistic effects are too large (Loucks 1967a) to be neglected in high Z materials. Relativistic band structure calculations are to be used in such cases. Now question arises as to how to compute the positron wavefunction in solids of high Z values. It is the purpose of this article to outline the various methods and to show the adequacy of the non-relativistic treatment of positrons even for solids of high Z value.

A naive calculation of the velocity of electron on the atomic model shows that the velocity ($V = Ze^2/n\nu$) increases with Z and attains the relativistic value soon. On the other hand a similar calculation of the positron velocity treating the positron in the Wigner-Seitz approximation in atoms of high Z value shows that the velocity is very low thereby not requiring a relativistic treatment. This justifies our findings that in positron annihilation studies in materials of high Z values, electrons are to be treated by relativistic methods whereas for positrons non-relativistic methods are quite sufficient. Different methods of treating positrons are outlined in section 2 along with a descriptions of the Wigner-Sietz method which we have used for our calculation. In Section 3 numerical results are presented along with a critical discussion of the results obtained.

2. Calculational methods

Various methods are followed for the calculations of positron wave function (Barko and Plaskett 1957, Stroud and Ehrenreich 1968, Gould et al 1972, Kubica and Stott 1974, Nieminen 1975). Assuming positron to be bound to an ion, one used the radial part of the wave function to be $r e^{-\alpha r}$ (A.T.Stewart, A.K.Pope, 1960) and α is used as a parameter. Sometimes the positron wave function is replaced by a δ function (Ferrell 1956). Later plane wave forms were assumed for positron. The plane wave for k = 0 state is

$$\Psi_+(r) = \Omega^{-\frac{1}{2}} \sum_g a_g e^{i g \cdot r} \tag{1}$$

and on substitution into the Schrodinger equation yields

$$\sum_{g'} \left[(g^2 - E)\, \delta_{g'g} + V_{g-g'} \right] a_{g'} = 0 \tag{2}$$

Here Ω represents the volume of the crystal,

$$V_{g-g'} = \int V(r)\, e^{i(g-g')\cdot r}\, d^3 r \tag{3}$$

and the expansion coefficients ag are normalized so that

$$\sum_g |a_g|^2 = 1 . \tag{4}$$

The high symmetry at k = 0 reduces the number of independent coefficients a_g thus facilitating the solution of the above equation. This method is useful in the case where one expects anisotropy in the region outside the ionic core. Pseudopotential

calculations are coming up in recent years (Kubica and Stott 1974, Nieminen 1975). The thermalized positrons preferentially occupy interstitial positions in the lattice as a consequence of the coulomb repulsion of the positive metallic ions. Wigner-Sietz method is found to be successful in describing the above situations. This method is employed (Rose and De Benedetti 1965, Terrell at al 1965, Loucks 1966, 1968, Chaung and Hogg 1967, Iyakutti et al 1976) for mostly in metals where one expects the positron wavefunction to be spherically symmetric. This becomes more suitable when one assumes the muffin-tin model potential for positron. This potential is computed in the same way as is done for electrons (Loucks 1967b), the differences being the change in the sign of the potential and the absence of exchange and correlation effects. In the computation of the muffin-tin potential for the positron the atomic charge densities of Herman and Skillman (1963) for Zr and those of Liberman (1970) for Th are used. The next step is to solve the Schrodinger equation or the Dirac equation with that potential using suitable boundary conditions. The radial part of the Schrodinger equation for positron with $\ell = 0$ can be written as

$$\frac{d^2 x(r)}{dr^2} + \left(E - V(r) \right) x(r) = 0 \qquad (5)$$

where V(r) is the potential seen by the positron. The solution $x(r) = r\,\psi_+(r)$ satisfies the following boundary conditions.

$$\psi_+(R) = 1 \text{ and } \psi_+'(R) = 0 \qquad (6)$$

where R is the radius of the APW sphere. For the relativistic treatment the radial part of the Dirac equation is to be used. The Dirac equations for positron (Rose 1961) are

$$\frac{df_+}{dr} = \frac{\eta - 1}{r} f_+ - (W - 1 + V) g_+ \qquad (7)$$

$$\frac{dg_+}{dr} = (W + V + 1) f_+ - \frac{\eta + 1}{r}$$

where f_+ and g_+ are respectively the small and large components, η = -1 and V(r) is the potential seen by the positron. It can be shown that the boundary condition for the solution of the Dirac equation reduces to f(R) = 0 and g(R) = 1 at the APW radius. With these boundary conditions the above equations are solved numerically

The most satisfactory method will be to treat the positron in the same way as electron and use the APW formalism. With the muffin-tin potential calculated for the positron the logarithmic derivative at the APW radius is computed. For k = 0 state, the zeros of the determinant are searched changing the energy. The lowest energy is taken as the appropriate eigenvalue for the positron and its wavefunction is determined for its corrsponding eigenstate.

3. Numerical Results and discussion

Numerical calculations are done first using the unscreened atomic potentials, $V(r) = Zq^2/r$. The mean square velocity of positron has also been calculated numerically, by finding the expectation value of the operator $V^2 = \frac{2}{m}(E - V)$. This was done for various value

of Z. It is found that the expectation value for the velocity of positron is very much less when compared to an electron in the same state for the same Z value (Table 1). In the atomic units employed in this article ($m = \frac{1}{2}$, $\hbar = 1$, $e^2 = 2.0$), the velocity of light $C \approx 274$. In fig.1, it is seen that the wave function obtained from the Schrodinger equation and the large component of the wave function computed from the Dirac equations coincide. The calculations are repeated with the muffin-tin potentials obtained for Zirconium (Z = 40) and Thorium (Z = 90) treating the positron both relativistically and non-relativistically (Figs.2 and 3). Once again a good agreement is obtained between the relativistic and non-relativistic calculations. The smaller component of the wave function is negligibly small compared to the large component. We have also done a detailed calculation of positron wave function using the APW formalism for Zr. Here as outlined earlier the lowest state corresponding to (k = 0) is taken (Stroud and Ehrenreich 1968, Gould 1972, Bross and Stohr 1974) and the contribution from l = 0 alone is taken into account. This result is compared with the wavefunction calculated numerically using the muffin-tin potential in the Schrodinger equation. The two results are shown graphically (Fig.2). It is seen from the graph that the values of the positron wave function obtained from APW formalism at various points between the centre and the boundary of the APW sphere are slightly less than the values obtained from the numerical integration of Schrodinger equation with muffin-tin-potential. It is not difficult to understand the origin of discrepancy. It may be pointed out that one of the boundary

conditions ($\psi_+ (R) = 0$) given in equation (6) is not satisfied in the APW formalism and as a consequence the eigenvalue (0.492495 Ryd) obtained in the APW formalism is different from the eigenvalue (0.792787 Ryd) which satisfies the boundary condition (6). However, if the eigenvalue obtained from the APW method is fed into the Schrodinger equation (5), the resulting solution will exactly coincide with the positron wave function obtained from the APW method.

The above investigation leads to the following conclusion. In the APW formalism the wave function of the positron in the eigenstate $\underset{\sim}{k} = 0$ does not differ much from the result obtained by solving the Schrodinger equation with the muffin-tin potential using the boundary conditions (6). This happens for any general direction and applicable for any polycrystalline sample. But in the case of oriented crystals, one has to use the APW formalism to retain the directional property of the positron wave function. Since the directional property of the positron wave function is exhibited only near the boundary of the cell (Stroud and Ehrenreich 1968, Bross and Stohr 1974, Singru 1974) and the value of the electron wave function almost approaches zero in that region, good results are expected even for oriented single crystals (Loucks 1966, 1968) even when the positron wave function simply obtained from numerical solution of Schrodinger equation with the muffin-tin potential is used. So even if one treats the electron wave function in the APW formalism, it is immaterial whether one treats the positron in the same footing or not for metals.

Inclusion of contribution from higher ℓ values in the APW calculation for the positron seems to have no effect on the eigenvalue obtained for positron. For instance in Zirconium the eigenvalue for the positron obtained is 0.492495 Ryd, restricting the ℓ value to zero whereas the inclusion of higher values upto $\ell = 6$ and $\ell = 10$ yields the eigenvalues 0.500365 Ryd and 0.500438 respectively. This is in contradiction to the case of electron for which the lowest eigenvalue obtained is 0.239347 Ryd whe $\ell = 0$ alone is included as against the eigenvalue 0.310468 obtained by including the ℓ values from 0 to 9.

Till now there has been no attempt on the relativistic treatment of positron annihilation in solids of high atomic number. From our calculations, it appears that one need not treat the positron wave function relativistically. This will mean that in the actual calculation of momentum density distribution in positron annihilation, the smaller component of electron wave function cannot contribute. It is because, in the final expression, there are two terms, one involving the product of large component of electron and positron wave functions and the other involving the product of their smaller components. It is interesting to note that even though there is no relativistic effect in the positron wave function the rate of positron annihilation is affected by the suppression of the contribution from the smaller component of electron wave function arising from the relativistic treatment of electron.

4. Conclusion

For high Z materials, one has to treat the electron relativistically but not neccssarily the positron in the positron annihilation studies. The most satisfactory method to obtain the positron wave function is to use the APW formalism but the less cumbersome procedure is to solve the Schrodinger equation with the muffin-tin potential with the boundary conditions specified in equation (6). The latter method however yields approximate but acceptable solutions.

Acknowledgements

We wish to thank Drs.T.Nagarajan, P.R.Subramanian and Mr.R.Asokamani for interesting discussion. The help in computation by Messrs. V.Aravamuthan, S.Sivasubramanian and V.Masilamani is acknowledged with thanks. We are thankful to Prof.D.Liberman for making available his atomic charge density data for thorium. Financial support from the University Grants Commission is also acknowledged with thanks.

References

Berko, S. and Plaskett, J.S., 1958, Phys. Rev. 112, 1877.

Bross, H. and Stohr, H., 1974, Appl. Phys. 3, 307 - 11.

Chuang, S.Y. and Hogg, B.G., 1967, Canad. J. Phys. 45, 3895.

Ferrell, R.A., 1956, Rev. of Modern Phys. 28.

Gould, A.G., West, R.N. and Hogg, B.G. 1972 Canad. J. Phys. 50 2294 - 301.

Herman, F. and Skillman, S. 1963, Atomic Structure Calculations (Englewood Cliffs, New Jersy: Prentice - Hall).

Iyakutti, K., Majumdar, C.K., Rao, R.S. and Devanathan, V. 1976 J. Phys., F. 6 (in press)

Kubica, P. and Stott, M.J., 1974, J. Phys. F. Metal Phys. 4 1969 - 81.

Liberman, D. 1970, Phys. Rev. B2, 244.

Loucks, T.L. 1966, Phys. Rev. 144, 504.

Loucks, T.L. 1967a Augmented Plane wave methods(New York : Benjamin) p.15

Loucks, T.L. 1967b Augmented Plane wave method (New York:Benjamin) pp 47 - 55.

Loucks, T.L. 1968 Phys. Rev. 176, 848

Nieminen, R.M. 1975, J. Phys. C. Solid State Phys. 8 2077 - 84

Rose, K.L. and DeBenedetti 1965 Phys. Rev. A138 927

Rose, M.E. 1961 Relativistic Electron Theory (New York, John Wiley and Sons) pp 157 - 61.

Singru, R.M. 1974 Pramana, 2 no 6 299-303.

Stroud, D. and Ehrenreich, H. 1968, Phys. Rev. 171, 399-407

Stewart, A.T., Pope A.K., 1960, Phys. Rev. 120, 15

Terrell, J.H., Weisberg and Berko, S. 1965 Positron annihilation Proceedings of the conference held at Wayne State University pp 269 - 276.

Table 1

Variation of electron and positron velocities with atomic number

Atomic number Z	Velocity (expectation value in a.u.)	
	Electron	Positron
10	20	2.217
20	40	2.934
30	60	3.411
40	80	3.778
50	100	4.085
60	120	4.358
70	140	4.611
80	160	4.852
90	180	5.082

Figure Legends

Fig.1: Comparison of the positron wave function obtained by solving the Schrodinger equation with the large component of the positron wave function computed from the Dirac equation using the atomic potential $V(r) = Ze^2/r$. Since both the functions coincide, a single curve is drawn for each element. The solid curve is obtained for Z = 40 and the dotted curve is for Z = 90. In both cases, the radius of the APW sphere is assumed to be 2.858 a.u.

Fig.2: Comparison of the positron wave function (solid curve) obtained by solving either the Schrodinger equation or the Dirac equation using the muffin-tin potential with the wave function (dotted curve) obtained by the APW formalism, for Zirconium (Z = 40).

Fig.3: Comparison of the positron wave function obtained by solving the Schrodinger equation with the large component of the positron wave function computed from the Dirac equation using the muffin-tin potential. A single curve is drawn since both the functions coincide. The APW sphere radius is R = 3.39694 a.u.

Fig.1.

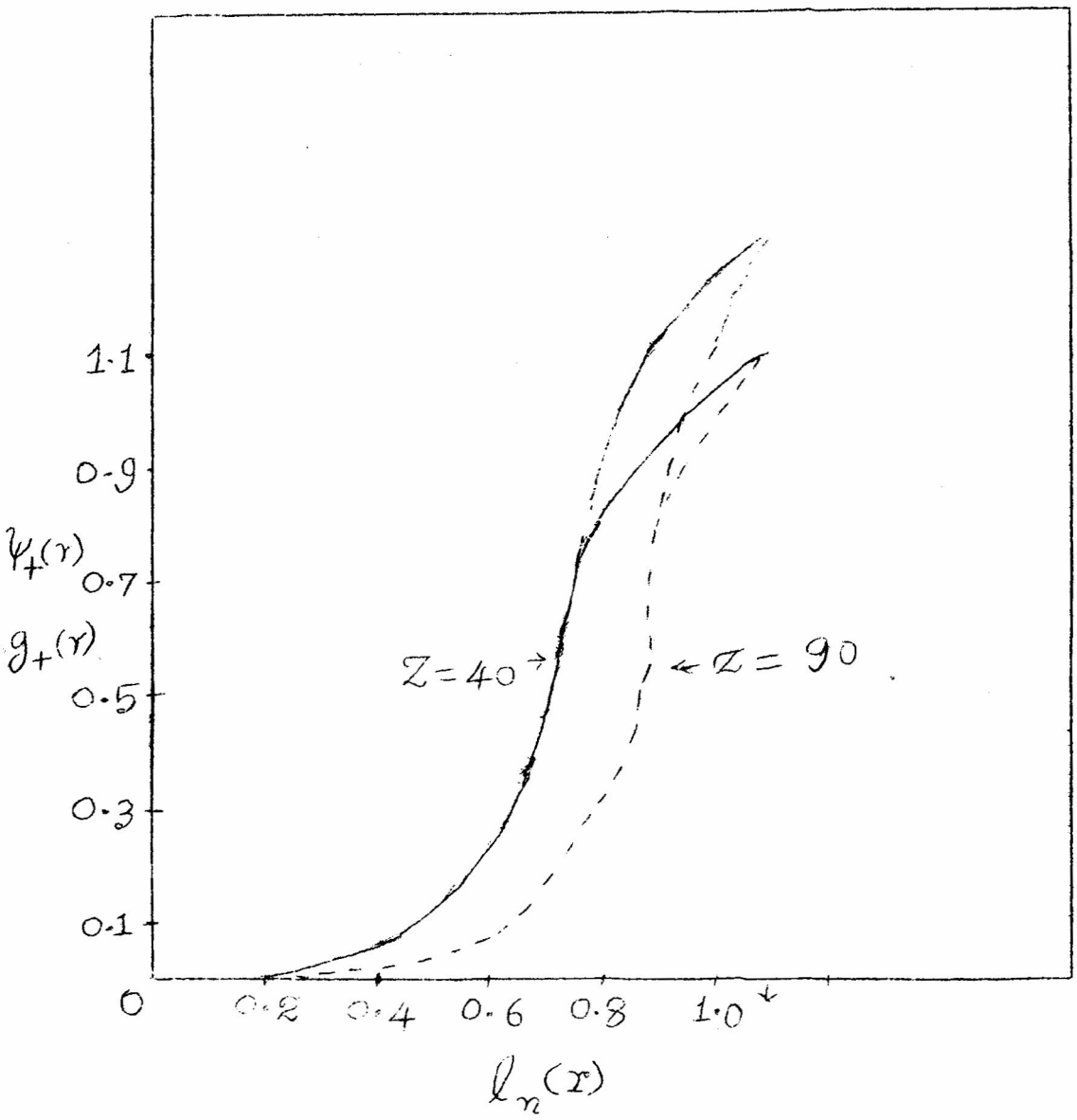


Fig.2

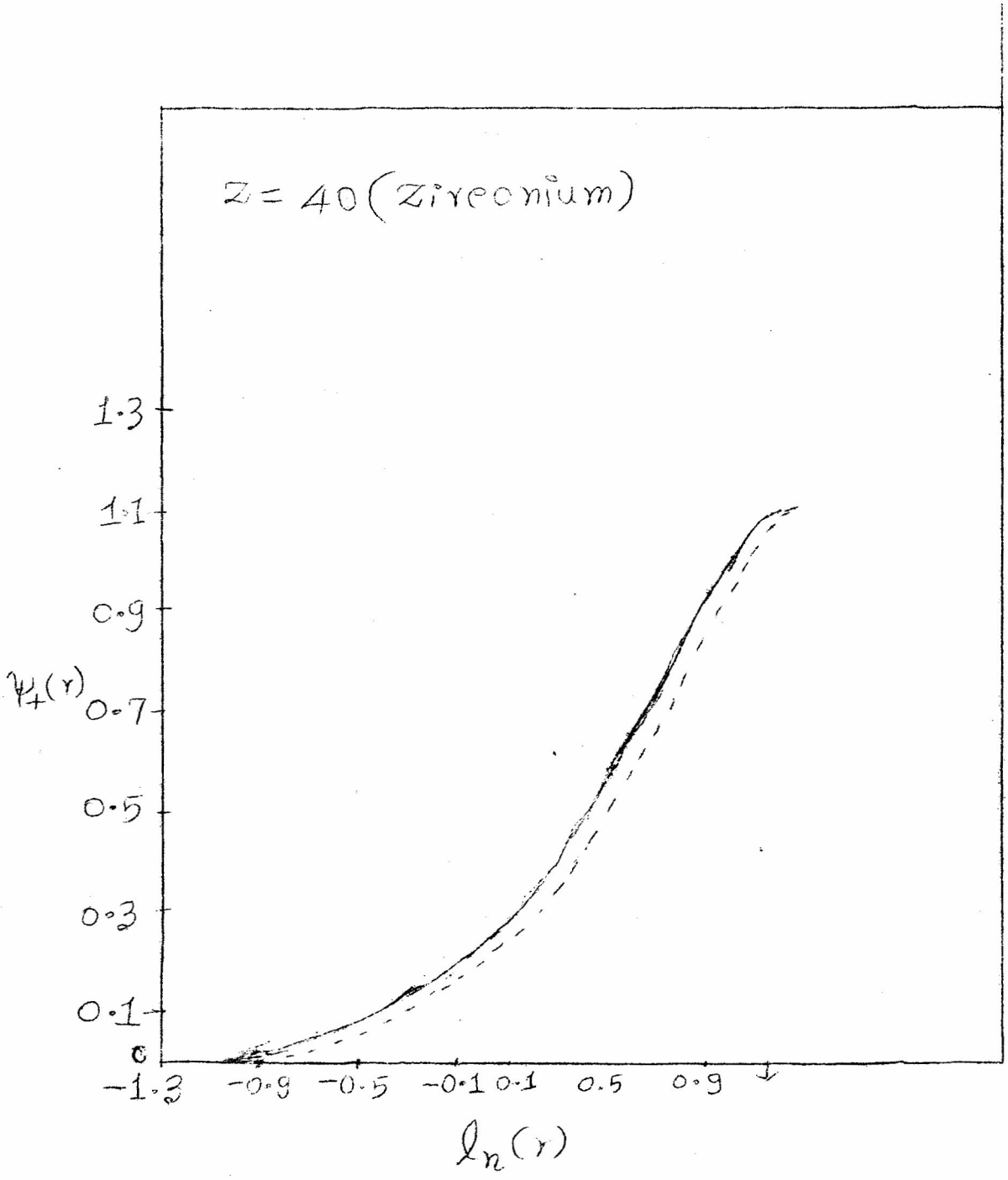
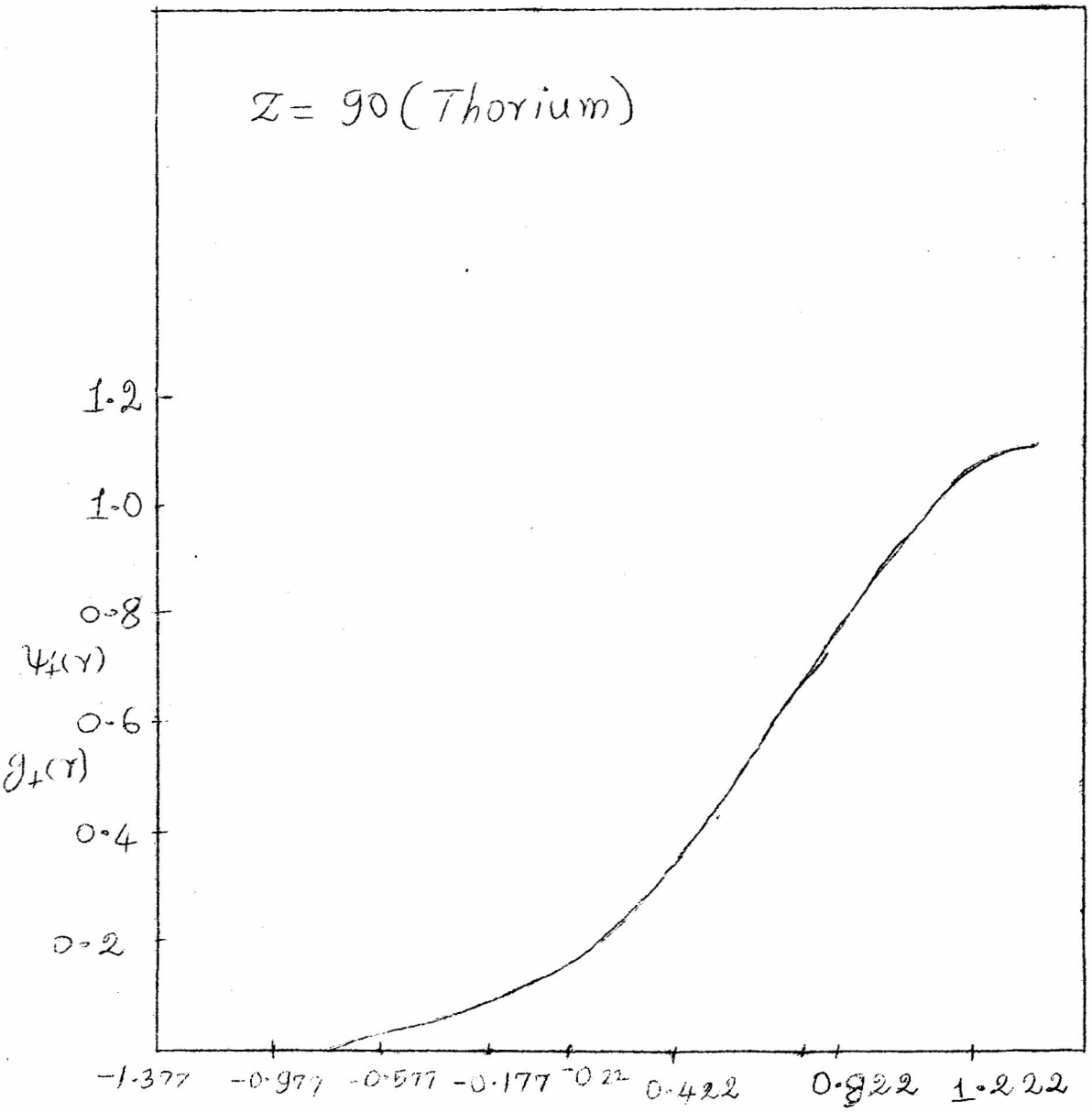


Fig.


Z = 90 (Thorium)
1·2
1·0
0·8
ψ+(γ)
0·6
g+(γ)
0·4
0·2
-1·377
-0·977
-0·577
-0·177
0·22
0·422
0·822
1·222