J. Phys. A: Math. Gen., Vol. 9, No. 8, 1976. Printed in Great Britain.

A simple proof of the Perron–Frobenius theorem for positive symmetric matrices

F Ninio

Department of Physics, Monash University, Clayton, Victoria 3168, Australia

Received 5 April 1976

Abstract. An elementary proof is given that the statistical mechanical transfer matrix, when symmetric, has a maximum eigenvalue which is non-degenerate and larger than the absolute value of any other eigenvalue. Moreover, the corresponding eigenvector can be chosen so that all its entries are strictly positive.

1. Introduction

In statistical mechanics the transfer matrix $\mathbf{A}$ allows the partition function of a system to be expressed in the form (Newell and Montroll 1953)

$$Z=\mathrm{Tr}(\mathbf{A}^N)=\sum_j \lambda_j^N$$

where the $\{\lambda_j\}$ are the eigenvalues of $\mathbf{A}$. From this one obtains the free energy F per particle in the thermodynamic limit as

$$F=-kT\lim_{N\to\infty} N^{-1}\ln Z=-kT\ln\lambda_{\max},$$

provided $\lambda_{\max}$ is positive, non-degenerate and greater than the absolute value of any other eigenvalue. These properties of the maximum eigenvalue are guaranteed for any square matrix $\mathbf{A}$ with elements $a_{ij}>0$ by the Perron–Frobenius theorem (Bellman 1970). There are many proofs of this result (see for instance the references in Bellman 1970), but all are quite long and by their generality tend to obscure the origin of the result. In most applications to statistical mechanics however one is concerned with a matrix which is symmetric ($a_{ij}=a_{ji}$) in addition to being positive ($a_{ij}>0$). A much simpler derivation can then be given. This is the object of the present paper.

2. Theorem and proof

Let $\mathbf{A}=(a_{ij})$ be an $n\times n$ symmetric matrix with elements $a_{ij}>0$ and let λ be the largest eigenvalue. Then

(i) $\lambda>0$

(ii) there exists a corresponding eigenvector (x_j) with every entry $x_j>0$

(iii) λ is non-degenerate

(iv) if μ is any other eigenvalue, $\lambda>|\mu|$.

Proof

(i) Since the eigenvalues of **A** are real and their sum equals Tr **A**> 0, it follows that $\lambda > 0$.

(ii) Let (u_j) be any real normalized eigenvector belonging to λ,

$$\lambda u_i = \sum_j a_{ij} u_j \qquad (i = 1, 2 \ldots n), \tag{1}$$

and set $x_j = |u_j|$. Then

$$0 < \lambda = \sum_{ij} a_{ij} u_i u_j = \left| \sum_{ij} a_{ij} u_i u_j \right| \leq \sum_{ij} a_{ij} x_i x_j.$$

By the variational theorem, the right-hand side is less than or equal to λ, with equality if and only if (x_j) is an eigenvector belonging to λ. We therefore have

$$\lambda x_i = \sum_j a_{ij} x_j \qquad (i = 1, 2 \ldots n). \tag{2}$$

Now if $x_i = 0$ for some i, then on account of $a_{ij} > 0$ for all j, it follows every $x_j = 0$, which cannot be. Thus every $x_i > 0$.

(iii) If λ is degenerate, we can find (since **A** is real symmetric) two real orthonormal eigenvectors (u_j), (v_j) belonging to λ. Suppose that $u_i < 0$ for some i. Adding equations (1) and (2), we get $0 = \lambda(u_i + |u_i|) = \Sigma_j a_{ij}(u_j + |u_j|)$, and as above, it follows $u_j + |u_j| = 0$ for every j. In other words we have either $u_j = |u_j| > 0$ for every j, or $u_j = -|u_j| < 0$ for every j. The same applies to (v_j). Hence $\Sigma_j v_j u_j = \pm \Sigma_j |v_j u_j| \neq 0$, i.e. u and v cannot be orthogonal. Thus λ is non-degenerate.

(iv) Let (ω_j) be a normalized eigenvector belonging to $\mu < \lambda$,

$$\sum_j a_{ij} \omega_j = \mu \omega_i.$$

The variational property and the non-degeneracy of λ now yield

$$\lambda > \sum_{ij} a_{ij} |\omega_i| |\omega_j| \geq \left| \sum_{ij} a_{ij} \omega_i^* \omega_j \right| = |\mu|.$$

Acknowledgments

I am indebted for Professor H Falk for bringing this problem to my notice, as well as for helpful correspondence on the subject. Part (iv) was communicated to Professor Falk by Professor F Wu.

References

Bellman R 1970 *Introduction to Matrix Analysis* (New York: McGraw-Hill)
Newell G F and Montroll E W 1953 *Rev. Mod. Phys.* **25** 353–89

J. Phys. A: Math. Gen., Vol. 10, No. 7, 1977. Printed in Great Britain. © 1977

Corrigenda

A simple proof of the Perron–Frobenius theorem for positive symmetric matrices
Ninio F 1976 *J. Phys. A: Math. Gen.* **9** 1281

Part (iv) of the proof should read as follows.

Let (ω_j) be a real normalized eigenvector belonging to $\mu < \lambda$,

$$\sum_j a_{ij}\omega_j = \mu\omega_i.$$

By the variational theorem,

$$\lambda \geqslant \sum_{ij} a_{ij}|\omega_i|\,|\omega_j| \geqslant \left|\sum_{ij} a_{ij}\omega_i\omega_j\right| = |\mu|.$$

If $\mu = -\lambda$, the above relation shows that $|\omega_j| = x_j$ for all j, and hence there is an i for which $\omega_i = x_i$. Adding $\lambda x_i = \Sigma_j a_{ij}x_j$ to $-\lambda\omega_i = \Sigma_j a_{ij}\omega_j$ gives

$$0 = \sum_j a_{ij}(x_j + \omega_j) \geqslant a_{ii}(x_i + \omega_i)$$

which contradicts the fact that $a_{ii} > 0$ and $\omega_i = x_i > 0$. Thus $\mu \neq -\lambda$.

The time structure of atmospheric Čerenkov light in extensive air showers
Böhm E, Bosia G, Navarra G and Saavedra O 1977 *J. Phys. A: Math. Gen.* **10** 441–60

The vertical axis of figure 2 (p 443) should read 'Anode current (arbitrary units)' and this same axis should not appear in figure 19 (p 459).

The caption for figure 15 (p 456) should read as follows.

Figure 15. Calculated energy spectrum of bursts compared with measured frequencies. Calculated frequencies of bursts: 1, protons; 2, iron primaries (where (a) refers to residual primaries and (b) to secondaries); 6, threshold burst energies. Measured fluxes: 3, bursts + Čerenkov light ($\theta = 1\cdot5°$); 4, bursts + Čerenkov light ($\theta = 4°$); 5, bursts only ($\theta = 20°$). θ is the assumed opening angle of the detector.

The fifth line of the second paragraph of appendix 3 (p 457) should read '. . . (see figure 3) . . . .

Also the equation at the bottom of p 458 should read:

$$n(\alpha) \approx \exp(-\alpha/\alpha_0)\,\mathrm{d}\omega\,\cos^n(\vec{\theta} + \vec{\alpha}).$$

32-vertex model on the triangular lattice
Sacco J E and Wu F Y 1975 *J. Phys. A: Math. Gen.* **8** 1780–7

The factors $2c$, $2d$, $2e$ in (21) should read $4c$, $4d$, $4e$; the definitions of a and b in (22) should be interchanged; the third line in (22) should read $\Omega_5\Omega_6 = f_{16}f_{34} + \bar{f}_{16}\bar{f}_{34} + f_0 f_{25} + \bar{f}_0\bar{f}_{25}$.

These changes do not alter any of the discussions and conclusions of the paper. We are indebted to K Y Lin and I P Wang for calling our attention to these corrections.