Part II

Theory and implementation

Foreword to Part II

Part II deals with the development and fundamental formulation of the theory of *Process Physics*, which represents a new paradigm – a self-referential and self-organizing bootstrap model that provides an alternative to the standard model of physics. This part is organized as follows:

- **Chapter 6** documents the theoretical and computational methods that provided the basis and founding insights for later progress. The work there should be regarded as a 'prototheory'.

- **Chapter 7** is a report that documents the theoretical and computational methods that provide the functional underpinning for *Process Physics*, from which principles all subsequent work in *Process Physics* derives.

- **Chapter 8** outlines and sketches the further developments that extend the foundations of *Process Physics* beyond its formative base, illustrating the functionality of the paradigm and its potential as a viable successor theory.

Chapter 6

Towards a new model: Heraclitean quantum systems

The cosmic order has been made by no god or man,
but has always been and always will be:
fire ever living, forever breaking out here and dying out there.

– Heraclitus

6.1 Introduction

In the standard model, one first constructs a classical spacetime. Next, various classical fields are attached to this geometrical structure. As a third stage, the matter fields are *quantized*. Eventually, with some hindsight, one might even seek to recover the classical behaviour of large quantum systems via some 'classicalization' argument or procedure. This multi-tiered approach reveals an incomplete separation of the historical (and necessarily contrived) development of the standard model, from a cohesive and comprehensive theoretical foundation. One would prefer to see, in a mature

theory, that the classical features of an objective reality emerge naturally from an abstract scheme that is itself devoid of classical structures and which is parsimonious in its assumptions and axioms.

The term, *Heraclitean quantum systems* (HQS), named for Heraclitus of Ephesus (540-480 BC), was coined and introduced in [343] as a basic model – a starting point, or precursor theory – that is constructed by first abandoning those features that one hopes will be emergent. A HQS is free of classical structures or concepts in its axioms. Observing that bosons are always composites of fermions, and that quark/antiquark pairs (i.e. fermions) obeying the Pauli Exclusion Principle (PEP) combine to form mesons, spacetime and all bosons are discarded in the assumption that these are low energy phenomena, leaving only fermions and the PEP. A HQS is a model that is constructed using a Grassmann algebraic system[1], which has algebraic fermion-like elements and anti-commutative rules that can be considered to play the rôle of an abstract form of the PEP.

6.2 Grassmann algebra and a correlation function

Let G_N be a Grassmanian algebra with N anti-commuting generators, m_N so that

$$\{m_i, m_j\} \equiv m_i m_j + m_j m_i = 0 \quad \forall \quad i, j = 1, \ldots, N \tag{6.2.1}$$

and so $(m_i)^2 = 0$. The 2^N polynomials:

$$1, m_1, \ldots, m_N, (m_1 m_2), (m_1 m_3), \ldots, (m_{N-1} m_N), \ldots, (m_1 m_2 \ldots m_N)$$

[1]Such algebras are used to very good effect in modelling the fermionic sector of the standard model. In such applications, however, multi-component local 'algebraic fields' are constructed and attached to some pre-specified spacetime manifold.

form a basis for G_N.

Similarly, a Grassmann algebra G_{2N} with $2N$ elements,

$$\overline{m}_1, m_1, \overline{m}_2 m_2, \ldots, \overline{m}_N m_N$$

(eventually with $N \to \infty$), may be defined, where $\overline{m}_i$ and m_i are completely independent (i.e. no relation between them is to be implied) and each pair of N generators anti-commutes with itself, so:

$$\{m_i, m_j\} \equiv 0, \quad \{\overline{m}_i, \overline{m}_j\} \equiv 0. \tag{6.2.2}$$

There is no distinction to be made between $\overline{m}_i$ and m_i until some form for the 'action', $S_{HQS}[\overline{m}, m]$ is specified. A 'correlation' function constructed using this Grassmann algebra is given by

$$G^{j,\ldots}_{i,\ldots} = \mathcal{G}\left[\overline{m}_i m_j \ldots \ \mathrm{e}^{-S_{\mathrm{HQS}}[\overline{m},m]}\right] \tag{6.2.3}$$

where the symbol $\mathcal{G}$ is used in the sense of Grassmann 'integration' over $\overline{m}, m$ 'fields':

$$\mathcal{G} \equiv \int \mathcal{D}\overline{q}\mathcal{D}q \ldots \tag{6.2.4}$$

where $\int \mathcal{D}X$ is used equivalently to $\int \prod_{i=1}^{n} dX_i$.

This '$\mathcal{G}$-*process*' is a purely algebraic device [344]. Specifically, the integral-like operation is defined by

$$\begin{aligned} \int dm_i \equiv & 0, \\ \int m_i \, dm_i \equiv & 1, \\ \{m_i, dm_j\} \equiv \{dm_i, m_j\} \equiv & 0 \quad \forall \quad i, j \in 1, \ldots N \end{aligned} \tag{6.2.5}$$

from which it follows that

$$\begin{aligned}
\int m_i\, dm_j + \int dm_i\, m_j =& \delta_{ij}, \\
\int m_i m_j\, dm_i\, dm_j = -& \int m_i \left(\int m_j\, dm_j \right) dm_i \qquad (i \neq j) \\
= -& \int m_i\, dm_i \quad = \quad -1
\end{aligned} \tag{6.2.6}$$

$$\begin{aligned}
\int m_i m_j\, dm_i =& - m_j \int m_i\, dm_i = -m_j \\
\int m_i\, dm_j dm_i =& - \int m_i\, dm_i dm_j = - \int dm_j = 0
\end{aligned} \tag{6.2.7}$$

and, in particular,

$$\begin{aligned}
\int m_{k_1} \ldots m_{k_N}\, dm_{k_N} \ldots dm_{k_1} =& \varepsilon_{k_1 \ldots k_N} \\
\int f(m)\, dm_{k_N} \ldots dm_{k_1} =& \sum_{k_1 \ldots k_N} \varepsilon_{k_1 \ldots k_N} f_{k_1 \ldots k_N}
\end{aligned} \tag{6.2.8}$$

where $\varepsilon_{k_1 \ldots k_N}$ is the antisymmetric tensor. It should be noted that in (6.2.8) only the N^{th}-order contribution to $f(m)$ survives the integration over the Grassmann generators.

Such primitive 'correlations' as (6.2.3) are necessarily nâive, being beneath any direct observation; they are not expected, in themselves, to have any phenomenological significance. A more general expression is given by

$$G[F] = \mathcal{G}[F[\overline{m}, m]\ \mathrm{e}^{-S_{HQS}[\overline{m}, m]}] \tag{6.2.9}$$

where F is some function of all the elements. To define the integration process,

expand the 'integrand' as a polynomial:

$$F[\overline{m}, m]e^{-S_{HQS}[\overline{m},m]} = 1 + \sum c_i m_i + \cdots + c_L \overline{m}_1 \overline{m}_2 \cdots \overline{m}_N m_1 m_2 ... m_N, \qquad (6.2.10)$$

writing the product of the highest order terms in some (arbitrary) standard order. By definition, and in a manner that will be demonstrated in what follows, only the coefficient of the 'last' term in the expansion survives the $\mathcal{G}$-process, so that:

$$\mathcal{G}[F[\overline{m}, m]\ \mathrm{e}^{-S_{HQS}[\overline{m},m]}] = c_L \qquad (6.2.11)$$

Consider now the particular HQS defined by the *quartic* action

$$S_{\mathrm{HQS}}\,[\overline{m}_i, m_j] = \sum_{i>j}^{N} \overline{m}_i m_j \overline{m}_j m_i \qquad (6.2.12)$$

(Henceforth, quotation marks are dropped from familiar terms borrowed for this exercise. The terminology adopted is used in an analogous sense rather than a literal one.) The exponential term in $G^{j,\cdots}_{i,\cdots}$ is replaced by the series expansion:

$$\mathrm{e}^{-S[\overline{m}_i, m_j]} = 1 - S + \frac{S^2}{2!} - \frac{S^3}{3!} + \cdots \qquad (6.2.13)$$

Clearly, the first term in the expansion contains no $\overline{m}, m$ pairs. The second term, S, is a sum of terms each containing two $\overline{m}, m$ pairs

$$\begin{aligned} S_{\mathrm{HQS}}\,[\overline{m}_i, m_j] = {} & \overline{m}_2 m_1 \overline{m}_1 m_2 + \overline{m}_3 m_1 \overline{m}_1 m_3 + \cdots + \overline{m}_N m_1 \overline{m}_1 m_N \\ & + \overline{m}_3 m_2 \overline{m}_2 m_3 + \overline{m}_4 m_2 \overline{m}_2 m_4 + \cdots + \overline{m}_N m_2 \overline{m}_2 m_N \\ & + \cdots \\ & + \overline{m}_N m_{N-1} \overline{m}_{N-1} m_N \end{aligned} \qquad (6.2.14)$$

The third term, S^2, must be a sum of terms each containing four $\overline{m}, m$ pairs of the form:

$$\overline{m}_i m_j \overline{m}_j m_i \overline{m}_{i'} m_{j'} \overline{m}_{j'} m_{i'}$$
$$\text{with} \quad i > j,\ i' > j',\ i \neq i',\ j \neq j',\ i \neq j',\ j \neq i'$$

(these conditions resulting as a consequence of the Grassmann algebra since any term with degeneracy in the indices is identically zero by definition). The nth term in the expansion is thus a sum consisting of terms each with $2(N-1)$ $\overline{m}, m$ pairs.

For a given number N of $\overline{m}, m$ pairs there will be some number k such that S^k is a sum of terms each with exactly N distinct $\overline{m}, m$ pairs if N is an even integer, or $N-1$ pairs if N is an odd integer. Then $S^{k+l} = 0$ for all l since these higher order terms are necessarily degenerate.

In fact, it will be seen that:

$$k = n_{\max} = \begin{cases} \frac{N}{2}, & N \text{ even} \\ \frac{N-1}{2}, & N \text{ odd} \end{cases} \tag{6.2.15}$$

To establish a numerical value for the non-zero components in these sums each 'string' of $\overline{m}, m$ pairs must be placed into some (arbitrary) standard order by pair-wise interchange of the m's. For each such exchange, the 'value' of the string is multiplied by a factor of (-1) so that an even number of interchanges to achieve the standard order results in an overall factor of $(+1)$, while an odd number of interchanges results in an overall factor of (-1); i.e. each string may be characterised by a *signature* of ± 1. If the standard order is chosen to be

$$\overline{m}_1 \overline{m}_2 \cdots \overline{m}_N m_1 m_2 \cdots m_N,$$

so that $\overline{m}_i m_j \overline{m}_j m_i, (i > j) \rightarrow \overline{m}_i \overline{m}_j m_j m_i \rightarrow \overline{m}_j \overline{m}_i m_j m_i \Rightarrow$ signature $= +1$, and similarly for each group of four ms, it is immediately obvious that all strings may be

rearranged thus with no change of signature:

$$\overline{m}_i m_j \overline{m}_j m_i \overline{m}_{i'} m_{j'} \overline{m}_{j'} m_{i'}, \quad (i > j,\ i' > j')$$

$$\rightarrow \quad \overline{m}_j \overline{m}_i m_j m_i \overline{m}_{j'} \overline{m}_{i'} m_{j'} m_{i'} \quad \Rightarrow \quad \text{signature } = +1$$

A further four interchanges collects the barred and unbarred ms giving:

$$\overline{m}_j \overline{m}_i \overline{m}_{j'} \overline{m}_{i'} m_j m_i m_{j'} m_{i'}$$

with no net change of signature. It remains now to order the indices within the barred and unbarred groups but notice that the indices are now arranged the same way for the $\overline{m}$s and the ms – any further rearrangement necessarily requires an even number of interchanges and so the signature for any non-zero string must be $(+1)$ for this choice of standard order. In fact, it is clear that this result must hold for any choice of standard order if we start with an action S of the form given.

Evaluation of S^n may now proceed, recognizing it to be equal to the number of non-degenerate terms in the sum. For $n = 1$, this is just the number of ways of choosing two distinct indices from the N available i.e. the binomial coefficient ${}_nC_2$. For $n = 2$, two more (different) indices are required – there are ${}_{(N-2)}C_2$ ways of choosing these and so the total number of non-degenerate terms in the sum is the product ${}_NC_2 \cdot {}_{(N-2)}C_2$. In general, write:

$$S^n = \prod_{i=0}^{n-1} {}_{(N-2i)}C_2, \quad n \leqslant \begin{cases} \frac{N}{2}, & N \text{ even} \\ \frac{N-1}{2}, & N \text{ odd} \end{cases} . \tag{6.2.16}$$

The condition on n is required otherwise the binomial coefficient is undefined for $(N - 2i) < 2$; this provides the $k = n_{\max}$ noted previously.

These results need to be incorporated in the overall calculation of the correlation function (6.2.3)

$$G^{i,\ldots}_{j,\ldots} = \mathcal{G}[\overline{m}_i m_j \ldots \ \mathrm{e}^{-S_{\mathrm{HQS}}[\overline{m},m]}]$$

Notice that for N even, S^k with $k = n_{\max}$ uses all N indices, leaving none for the leading $\overline{m}_i m_j$ term(s) whereas for N odd, S^k uses $N-1$ indices. Thus, for N even the highest order term in the expansion is given by S^{k-1}, leaving two indices free – this demands that the leading term in G take the form $\overline{m}_i m_j \overline{m}_j m_i$ to ensure that each index is used and symmetry is maintained in $\overline{m}m$ pairs. On the other hand, for N odd the highest order term in the expansion leaves just one index free and the leading term in G has the form $\overline{m}_i m_j$, with $i = j$. Hence one is obliged to consider at least two distinct classes of solution for G for this particular choice of the action, $S[\overline{m}_i, m_j]$: those with N even and those with N odd.

Extending the Grassmann algebra to include sources $\bar{l}_i, l_i$, a generating functional is introduced

$$\begin{aligned} Z\left[\bar{l}, l\right] &= \mathcal{G}\left[\exp(-S_{\text{HQS}}\left[\overline{m}, m\right] - \bar{l}m - \overline{m}l\right] \\ &= \mathcal{G}\left[\exp\left(-\sum_{i>j} \overline{m}_i m_j \overline{m}_j m_i - \sum_i \left(\bar{l}_i m_i + \overline{m}_i l_i\right)\right)\right] \\ &= 1 + l_i G_i + \bar{l}_i G_i + G_{ij} \bar{l}_i l_j + \cdots \end{aligned} \tag{6.2.17}$$

No notion of locality is permissible, so all elements are in 'interaction'. The apparent dominance of local interactions must be emergent. The action has a large invariance group: $m \to Um, \overline{m} \to \overline{m}U^{-1}$. To establish the methodology to be employed next, bosonization of the generating functional (6.2.17), it is first pertinent to reflect on what has been learned elsewhere.

6.3 Action sequencing

Quantum field theory achieves extraordinary accuracy in its predictions but the most fundamental model is that of Quantum Chromodynamics (QCD), which provides an example of the emergence of a complex effective theory. In QCD, the *functional integral calculus* (FIC) is employed to take advantage of the functional integral formulation of quantum field theories. Insights gained from the derivation [345, 346] of emergent hadronic phenomena for the quark-gluon system, in which the notion of *induced action sequencing* is highlighted as an essential aspect, provide a convenient introduction to the ideas to be developed here.

To illustrate the concept of action sequencing, observe the (inexact) derivation of the low-energy, or hadronic, form of QCD, beginning with the approximation of QCD by the Global Colour Model (GCM) [345, 346], *viz*:

$$\langle 0 \mid 0 \rangle_J = \int \mathcal{D}\overline{q}\,\mathcal{D}q\,\mathcal{D}A\, \exp\left(-S_{\rm QCD}\left[A,\overline{q},q\right]\right) \tag{6.3.1}$$

$$\approx \int \mathcal{D}\overline{q}\,\mathcal{D}q\,\mathcal{D}A\, \exp\left(-S_{\rm GCM}\left[A,\overline{q},q\right]\right) \quad (\textit{Global Colour Model}) \tag{6.3.2}$$

$$= \int \mathcal{D}\mathcal{B}\,\mathcal{D}\mathcal{D}\,\mathcal{D}\mathcal{D}^{\star}\, \exp\left(-S_{\rm bilocal}\left[\mathcal{B},\mathcal{D},\mathcal{D}^{\star}\right]\right) \quad (\textit{bilocal fields}) \tag{6.3.3}$$

$$= \int \mathcal{D}\pi \ldots \mathcal{D}\overline{N}\,\mathcal{D}N \ldots \exp\left(-S_{\rm hadronic}\left[\pi,\ldots,\overline{N},N,\ldots\right]\right) \quad (\textit{local fields}) \tag{6.3.4}$$

Each change of functional integration field variables (mandated by the dynamics) generates a new effective action for those field variables, so that:

$$S_{\rm QCD}\left[A,\overline{q},q\right] \to S_{\rm GCM}\left[A,\overline{q},q\right] \to S_{\rm\ bilocal}[\mathcal{B},\mathcal{D},\mathcal{D}^{*}] \to S_{\rm hadronic}[\pi,\ldots,\overline{N},N,\ldots] \tag{6.3.5}$$

A vital intermediate step is to determine the minimum of the action in (6.3.3):

$$\frac{\delta S_{\rm bilocal}[\mathcal{B},\mathcal{D},\mathcal{D}^{*}]}{\delta\mathcal{B}} = 0, \ \ldots \tag{6.3.6}$$

has a solution with $\overline{\mathcal{B}} \neq 0$, giving rise to the *condensate* effect in $\overline{q}q$ whereby the induced effective action has a non-trivial minimum away from the $\mathcal{B} = 0$ perturbative point. This is seen to be one of the most important dynamical effects in QCD in terms of comprehending the nature of hadrons and is similar to effects observed in superconductivity. If hadrons are considered as deviations in the structure of the condensate then the lowest mass hadrons correspond to those deviations that occur in the flattest directions of the effective action for bilocal fields in (6.3.3) (see [345] – [351]). These correspond to the Nambu-Goldstone (NG) pion modes, $U(x) = \exp(i\sqrt{2}\,\pi^a(x)F^a)$ with NG boson fields $\pi(x)$ forming homogeneous Riemann coordinates for the condensate manifold. The internal structure of pions is deeply related to the structure of the condensate and so the near-degeneracy of the condensate equations (6.3.6) determines the long-range behaviour of the nuclear force.

Let $\overline{\mathcal{B}}(x,y)$ be a solution of (6.3.6), then

$$\mathcal{B}(x,y) = \overline{\mathcal{B}}(x,y) + \sum_a \phi_a\left(\frac{x+y}{2}\right)\Gamma^a(x-y), \tag{6.3.7}$$

provides the idea of a deviation from the condensate. Here, the $x-y$ dependence of $\mathcal{B}(x,y)$ is expanded into a complete (and, in principal, arbitrary) set of functions, $\Gamma^a(w)$, with $\phi_a(z)$ as expansion coefficients. The integration variables may be changed from the $\mathcal{B}(x,y)$ to the $\phi_a(z)$. In the GCM, the $\Gamma^a(w)$ are chosen so as to diagonalize the second-order terms that arise when the bilocal effective action in (6.3.3) is expanded about $\overline{\mathcal{B}}(x,y)$. Consequently, the $\Gamma^a(w)$ are found to be solutions of Beth-Salpeter equations, describing the internal structure of $\overline{q}q$ core states. The 'centre-of-mass' motion of these mesonic bound states is described by the $\phi_a(z)$ fields, such as are found to occur in (6.3.4).

6.4 Bosonization

Returning to (6.2.17), to consider a bosonization along the lines of the GCM bosonization in QCD, look again at the generating functional:

$$\begin{aligned} Z\left[\bar{l},l\right] &= \mathcal{G}\left[\exp(-S_{\text{HQS}}\left[\overline{m},m\right]-\bar{l}m-\overline{m}l\right] \\ &= \mathcal{G}\left[\exp\left(-\sum_{i>j}\overline{m}_i m_j \overline{m}_j m_i - \sum_i \left(\bar{j}_i m_i + \overline{m}_i j\right)\right)\right] \end{aligned} \tag{6.4.1}$$

This use of the Grassmann algebra provides an analogue of the PEP by excluding $\overline{m}, m$ combinations with repeated indices – the condition $i > j$ in the first sum avoids both double counting and the redundant zero-value terms that would otherwise occur when $i = j$. It is found, however, that this fermionic structure is difficult to work with, thus the need to perform a 'bosonization'. First, it is shown that Z can be put in the form

$$Z\left[\bar{l},l\right] = \mathcal{G}\left[\int \mathcal{D}B \exp\left(-S_{Bmm}\left[B,\overline{m},m\right]-\bar{l}m-\overline{m}l\right)\right],$$

$$\text{where } S_{Bmm} = \tfrac{1}{2}\sum_{i,j} B_{ij}B_{ij} - \sum_{i,j} B_{ij}\left(\overline{m}_i m_j - \overline{m}_j m_i\right), \tag{6.4.2}$$

$$\text{with } B_{ij} = -B_{ji}, \quad B_{ii} = 0, \quad \text{and } B_{ij} \text{ real.}$$

This is checked by performing the Grassmann integrations; for convenience of notation, define

$$B^2 = \sum_{i,j} B_{ij}B_{ij} \text{ and } \Psi = \Psi_{ij} \equiv \left(\overline{m}_i m_j - \overline{m}_j m_i\right), \text{ with } \Psi_{ij} = -\Psi_{ji} \tag{6.4.3}$$

and drop the summation signs and indices, e.g. $\bar{l}m \Rightarrow \sum_i \bar{l}_i m_i$ etc, then write Z in (6.4.2) in the form

$$Z\left[\bar{l},l\right] = \iint \mathcal{D}\overline{m}\,\mathcal{D}m \exp\left(-\bar{l}m - \overline{m}l\right) \int \mathcal{D}B \exp\left(-\tfrac{1}{2}B^2 + B\Psi\right) \tag{6.4.4}$$

and perform the integration over B. To do this, complete the square

$$\begin{aligned} \tfrac{1}{2}B^2 - B\Psi &= \tfrac{1}{2}\left(B^2 - 2B\Psi\right) \\ &= \tfrac{1}{2}\left(B-\Psi\right)^2 - \tfrac{1}{2}\Psi^2 \end{aligned} \tag{6.4.5}$$

and make a change of variables to $Y_{ij} = (B - \Psi)$. Because the Ψ_{ij} are even-order Grassmann elements, they are mutually commuting and behave like real and complex numbers (often called c-numbers) so the usual change of variable rules still apply. Therefore

$$\int \mathcal{D}B \exp\left(-\tfrac{1}{2}B^2 + B\Psi\right) \quad \rightarrow \quad \int \mathcal{D}Y \exp\left(-\tfrac{1}{2}Y^2 - \tfrac{1}{2}\Psi^2\right). \tag{6.4.6}$$

Now, formally, $\int \mathcal{D}Y\, \mathrm{e}^{-\frac{1}{2}Y^2}$ is a Gaussian integral and thus equal to a constant and so

$$\begin{aligned} \int \mathcal{D}Y \exp\left(-\tfrac{1}{2}Y^2 - \tfrac{1}{2}\Psi^2\right) &= C\ \mathrm{e}^{-\frac{1}{2}\Psi^2} \\ \therefore Z\left[\bar{l},l\right] &\propto \int \mathcal{D}\overline{m}\,\mathcal{D}m \exp\left(-\tfrac{1}{2}\Psi^2 - \bar{l}m - \overline{m}l\right) \\ &= \int \mathcal{D}\overline{m}\,\mathcal{D}m \exp\left(-\tfrac{1}{2}\left(\overline{m}_i m_j - \overline{m}_j m_i\right)^2 - \bar{l}m - \overline{m}l\right) \end{aligned} \tag{6.4.7}$$

Expanding:

$$\tfrac{1}{2}\left(\overline{m}_i m_j - \overline{m}_j m_i\right)^2 = \tfrac{1}{2}\left(\overline{m}_i m_j \overline{m}_i m_j + \overline{m}_j m_i \overline{m}_j m_i - \overline{m}_i m_j \overline{m}_j m_i - \overline{m}_j m_i \overline{m}_i m_j\right) \tag{6.4.8}$$

– the first two terms on the RHS are identically zero (Grassmann property) and the Grassmann algebra provides that $\overline{m}_j m_i \overline{m}_i m_j = (-1)^4 \overline{m}_i m_j \overline{m}_j m_i$. Thus

$$Z\left[\bar{l}, l\right] \propto \iint \mathcal{D}\overline{m}\,\mathcal{D}m\ \mathrm{e}^{-\overline{m}_i m_j \overline{m}_j m_i - \bar{l}m - \overline{m}l}, \tag{6.4.9}$$

which demonstrates the equivalence of (6.2.17 (6.4.1)) and (6.4.4). Returning to the latter to expand in Ψ and explicitly perform the $\mathcal{G}$ process via the $\overline{m}, m$ integration:,

$$\begin{aligned} Z\left[\overline{m}, m\right] &= \iiint \mathcal{D}B\,\mathcal{D}\overline{m}\,\mathcal{D}m \exp\left[-\tfrac{1}{2}B^2 + \overline{m}Bm - \bar{l}m - \overline{m}l\right] \\ &= \int \mathcal{D}B\ \mathrm{e}^{-\frac{1}{2}B^2} \iint \mathcal{D}\overline{m}\mathcal{D}m\ \mathrm{e}^{\overline{m}Bm - \left(\bar{l}m + \overline{m}l\right)} \end{aligned} \tag{6.4.10}$$

Completing the square

$$\overline{m}Bm - \left(\bar{l}m + \overline{m}l\right) = \left(\overline{m} - \bar{l}\,B^{-1}\right) B \left(m - B^{-1}\,l\right) - \bar{l}\,B^{-1}l \tag{6.4.11}$$

and changing variables, $M = m - B^{-1}\,l$, to eliminate the linear term then

$$\overline{m}Bm - \left(\bar{l}m + \overline{m}l\right) = \overline{M}BM - \bar{l}\,B^{-1}l. \tag{6.4.12}$$

This change of Grassmann variables thus permits that

$$I = \iint \mathcal{D}\overline{m}\,\mathcal{D}m\ \mathrm{e}^{-\overline{m}Bm + \left(\bar{l}m + \overline{m}l\right)} = \iint \mathcal{D}\overline{M}\,\mathcal{D}M\ \mathrm{e}^{-\overline{M}BM + \bar{l}\,B^{-1}\,l}. \tag{6.4.13}$$

Now, the matrix exponential $\mathrm{e}^{A+B} = \mathrm{e}^A \mathrm{e}^B$ only if A and B commute, so consider the commutator

$$\begin{aligned} \left[-\overline{M}BM,\ \bar{l}\,B^{-1}\,l\right] &= \left(-\overline{M}BM\right)\left(\bar{l}\,B^{-1}\,l\right) - \left(\bar{l}\,B^{-1}\,l\right)\left(-\overline{M}BM\right) \\ &= \left(\overline{M}BM\right)\left(\bar{l}\,B^{-1}\,l\right) - \left(\bar{l}\,B^{-1}\,l\right)\left(\overline{M}BM\right) \\ &= \overline{M}M\,\bar{l}\,l - \bar{l}\,l\,\overline{M}M. \end{aligned} \tag{6.4.14}$$

Rearranging the second term by the usual pair-wise interchange of Grassmann elements provides the required commutation relation

$$\begin{aligned} \left[-\overline{M}BM,\ \bar{l}\,B^{-1}\,l\right] &= \overline{M}M\,\bar{l}\,l - (-1)^4\,\overline{M}M\,\bar{l}\,l \\ &= 0 \end{aligned} \tag{6.4.15}$$

so

$$\begin{aligned} I &= \iint \mathcal{D}\overline{M}\,\mathcal{D}M\ \mathrm{e}^{-\overline{M}BM}\ \mathrm{e}^{\bar{l}(B^{-1})\,l} \\ &= \mathrm{e}^{\bar{l}(B^{-1})\,l} \iint \mathcal{D}\overline{M}\,\mathcal{D}M\ \mathrm{e}^{-\overline{M}BM}. \end{aligned} \tag{6.4.16}$$

But $\int \mathcal{D}M\ \mathrm{e}^{-\overline{M}BM} = \det B$ and $\det B = \exp\left[\mathrm{TrLn}(B)\right]$ for any non-singular matrix B. Therefore, (6.2.17) can finally be written in the form

$$Z\left[\bar{l},\, l\right] = \int \mathcal{D}B \exp\left[-\sum_{i>j} B_{ij}^2 + \mathrm{TrLn}(B) + \bar{l}\, B^{-1}\, l\right]. \tag{6.4.17}$$

The Grassmann algebraic aspects are now contained in TrLn and the $\bar{l}\,l$ algebra. Notably, Z is independent of $\overline{m}, m$ terms and thus is non-fermionic (except for the source terms) – that is, the generating function has been 'bosonized'. (Note that the 'm-squared' form was used in (6.2.17) because only this form may be exactly bosonized. This arises because multi-dimensional Gaussian integrals can be evaluated exactly. Also, note that the '+' sign for the trace-log term arises from the properties of the $\overline{m}, m$s that were chosen so as to emulate the PEP). The algebraic $\mathcal{G}$ process has been given a representation that involves the sum over all possible B configurations. This is the Heraclitean 'flux', which is an axiomatic aspect of HQS. What is sought is the emergence of stability and 'things' – or, as Wheeler calls it, "It from Bit"[242, 172].

The expression for Z in (6.4.17) is exactly equivalent to (6.2.17) – however, it has the advantage of not being algebraic and one could imagine estimating (6.4.17) by appropriate approximations.

Taking the action in the form obtained in (6.4.17) and dropping the source terms gives the *effective* or *induced* action

$$S_{\mathrm{C}}\left[B\right] = \sum_{i>j} B_{ij}^2 - \mathrm{TrLn}(B), \tag{6.4.18}$$

where the label 'C' given to the action identifies it with the notion of a 'condensate' or 'ground-state' of the HQS. Here, 'B' is the analog of bilocal fields in QCD. The advantage of the bosonization technique is that it renders the system more tractable and better prepared for the very effective analytical methods practiced in QCD (particularly the GCM). Also, present day modelling has indicated the extraordinary success of mean field or smoothing approximations and it would be expected that these should be accessible via the bosonization – although it should be noted (as again illustrated by the GCM) that bosonization does not preclude the emergence of complex fermionic components.

The bosonization technique continues by finding and then analyzing the minimum of the induced action to identify the most significant part of the B integrations. To solve

$$\frac{\delta S_{\mathrm{C}}\left[B\right]}{\delta B} = 0, \tag{6.4.19}$$

let $B = \overline{B} + b$, where b is a small perturbation of the 'ground' state $\overline{B}$ such that

$$b_{ij} = -b_{ji}, \quad b_{ii} = 0 \quad \text{(antisymmetric)} \tag{6.4.20}$$

then, for all small b,

$$\begin{aligned} S\left[\overline{B}+b\right] - S\left[\overline{B}\right] &= \sum_{i>j}(\overline{B}_{ij}+b_{ij})^2 - \mathrm{TrLn}(\overline{B}+b) - \sum_{i>j}\overline{B}_{ij}^2 + \mathrm{TrLn}(\overline{B}) \\ &= \sum_{i>j}(\overline{B}_{ij}^2 + 2\overline{B}_{ij}b_{ij} + b_{ij}^2) - \sum_{i>j}\overline{B}_{ij}^2 - \mathrm{TrLn}(\overline{B}+b) + \mathrm{TrLn}(\overline{B}) \end{aligned} \tag{6.4.21}$$

$$\Rightarrow \quad S\left[\overline{B}+b\right] - S\left[\overline{B}\right] = \sum_{i>j} 2\overline{B}_{ij}b_{ij} - \mathrm{TrLn}(\overline{B}+b) + \mathrm{TrLn}(\overline{B}) + \mathrm{O}(b^2).$$

Now,

$$\begin{aligned}\mathrm{TrLn}(\overline{B}+b)-\mathrm{TrLn}(\overline{B}) &= \mathrm{TrLn}\left[\overline{B}(I+\overline{B}^{-1}b)\right]-\mathrm{TrLn}(\overline{B}) \\ &= \mathrm{TrLn}(\overline{B})+\mathrm{TrLn}(I+\overline{B}^{-1}b)-\mathrm{TrLn}(\overline{B}) \\ &= \mathrm{TrLn}(I+\overline{B}^{-1}b) \\ &= \mathrm{Tr}(\overline{B}^{-1}b)+\ \mathrm{O}(b^2),\end{aligned} \tag{6.4.22}$$

using the facts that $\ln(1+y)\approx y,\quad y\ll 1$ and $\overline{B}^{-1}b\ll I$, $I=N\times N$ identity matrix. Substituting the results from (6.4.22) in (6.4) and dropping $\mathrm{O}(b^2)$ terms (which is acceptable since b is small), at a minimum the following relation holds for all b_{ij}

$$\begin{aligned}\sum_{i>j}2\overline{B}_{ij}b_{ij} &= \mathrm{Tr}(\overline{B}^{-1}b) \\ &= \sum_{k,l}\overline{B}^{-1}_{lk}b_{kl} \\ &= -2\sum_{i>j}\overline{B}^{-1}_{ij}b_{ij}.\end{aligned} \tag{6.4.23}$$

Comparing the coefficients of b_{ij} thus provides the 'vacuum' equation

$$\begin{aligned}\overline{B}_{ij} &= \overline{B}^{-1}_{ij} \\ \text{i.e.}\quad \overline{B} &= \overline{B}^{-1}\end{aligned} \tag{6.4.24}$$

which is analogous to the Gap Equation of superconductivity BCS theory and also to chiral symmetry-breaking in QCD. Multiplying both sides by $\overline{B}$

$$\Rightarrow\quad \overline{B}\cdot\overline{B}=-I. \tag{6.4.25}$$

One is prompted to ask what form $\overline{B}$ must take to satisfy this relation. Notice that the 2×2 identity matrix and the 2×2 Pauli matrices $\sigma_x,\ \sigma_y,\ \sigma_z$ have the property

$\sigma_i^2 = I \quad (i = x, y, z)$ and so does $\hat{n} \cdot \vec{\sigma}$ ($\hat{n}$ is a unit 4-vector and $\vec{\sigma} = [\sigma_x, \sigma_y, \sigma_z, I]$), thus

$$(i\hat{n} \cdot \sigma)^2 = -I. \tag{6.4.26}$$

But σ_y is antisymmetric, which motivates one solution for $\overline{B}$ – the block-diagonal matrix of the form

$$\overline{B} = \begin{bmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & 0 & \cdots \\ 0 & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{6.4.27}$$

for even-order i, j. The general solution obtains by performing a similarity transformation on $\overline{B}$ so that $\overline{B} = R\overline{B}_0 R^{-1}$, where R is an arbitrary real orthogonal matrix and $\overline{B}_0$ is the block diagonal matrix given in (6.4.27). Clearly, this is highly degenerate (dominating the B-fluctuations) and contains none of the structure that is indicative of the usual meaning of the term, 'condensate'. The R-transformation 'switches' element pairings and the non-zero elements of each 2×2 block on the diagonal may be regarded in the sense of Cooper pairs, for example B_{12} couples $\overline{b}_1, b_2$ etc.

6.5 Condensate excitations

Consider the '*nihilo* $\rightarrow$ *nihilo* amplitude'

$$\langle \mathcal{N} \mid \mathcal{N} \rangle = Z[0] = \int \mathcal{D}B \; \mathrm{e}^{-S_{\mathrm{C}}[B]}, \tag{6.5.1}$$

and deviations from $\overline{B}_0$

$$B_{ij} = \overline{B}_{0_{ij}} + \sum_a \phi_a \Gamma^a_{ij}. \tag{6.5.2}$$

in which the Γ^a form an arbitrary complete set. The problem is to find a set that facilitates access to emergent phenomena (again, following the example of the GCM).

As a first choice, set a to be a serial index $a \equiv (IJ)$ and let

$$\Gamma^a_{ij} = \begin{cases} -\Gamma^a_{ji} = 1, & \text{if } I = i \text{ and } J = j \\ 0 & \text{otherwise} \end{cases} \tag{6.5.3}$$

These Γ^a_{ij} form a complete set for the expansion (6.5.2), with expansion coefficients ϕ_a. Then, changing variables of integration from B to ϕ

$$\langle \mathcal{N} \mid \mathcal{N} \rangle = \int \mathcal{D}\phi \; \mathrm{e}^{-S_{\mathrm{C}}\left[\overline{B}_0 + \phi \cdot \Gamma\right]} \tag{6.5.4}$$

which, in essence, is equivalent to the defining B_{ij} integrations (the Jacobian is a constant and can be ignored). It is usual, when considering degenerate condensates, to make the *superselection* assumption that it is legitimate to work in the neighbourhood of one condensate point. Choosing $\overline{B}_0$ and expanding S_{C} in powers of ϕ_a gives

$$S_{\mathrm{C}}\left[\overline{B}_0 + \phi \cdot \Gamma\right] = S_{\mathrm{C}}\left[0\right] + \sum_{ab} \phi_a \phi_b K_{ab} + \sum_{abc} \phi_a \phi_b \phi_c K_{abc} + \cdots \tag{6.5.5}$$

These variables of integration affect only small numbers of $\overline{m}, m$ pairings. They are too primitive to be able to reveal any complex emergent behaviour. Nevertheless, they may be partly analysed by diagonalising the quadratic term in (6.5.5) and thus choosing new variables of integration so that

$$\langle \mathcal{N} \mid \mathcal{N} \rangle = \int \mathcal{D}\Phi \; \mathrm{e}^{-S_{\mathrm{C}}[0] - \sum_a \Phi_a \Phi_a \lambda_a - \sum_{abc} \Phi_a \Phi_b \Phi_c K'_{abc} + \cdots} \tag{6.5.6}$$

This change of variables is equivalent to a new choice for the Γ^a_{ij}. Approximately one half of the eigenvalues λ_a have value zero: these correspond to the 'massless' NG modes, i.e. deviations in the tangent plane to the condensate manifold. The remaining λ_a are all non-zero and equal: these 'massive' modes correspond to deviations perpendicular to the condensate manifold (figure 6.1 illustrates the idea).

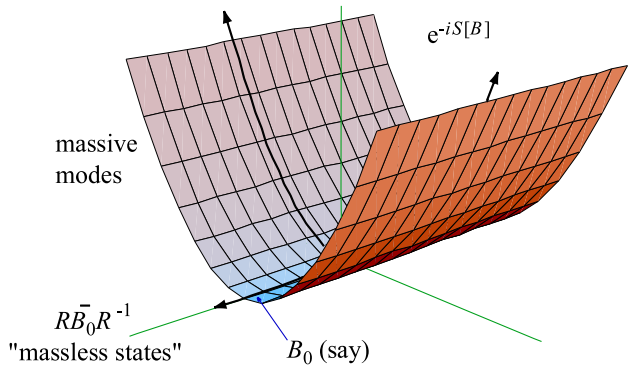


Figure 6.1: Condensate deviations
showing 'massless' modes in the direction of the tangent plane and 'massive' modes perpendicular to the manifold.

In QCD the analogue of the Φ_a modes are $\overline{q}q$ meson core-state modes, and the diagonalisation procedure is there the Bethe-Salpeter equation. Because of the peculiarities of QCD the hadrons contain either two constituent quarks (mesons) or three constituent quarks (baryons) together with secondary mesonic dressings of these core states. In HQS, however, multi-constituent modes are needed, since it is only with these that one might hope to find evidence of classical structures – and for that purpose the preceding two choices of integration variables are not helpful.

Consider now a third choice of integration variables, amounting to another selection for the complete set of Γ^a so that

$$\langle \mathcal{N} \mid \mathcal{N} \rangle = \int \mathcal{D}G \; \mathrm{e}^{-S_{\mathrm{C}}\left[\overline{B}_0 + G\cdot\Gamma\right]} \tag{6.5.7}$$

where

$$\Gamma = \{\Gamma^a : \Gamma^a_{ij} = -\Gamma^a_{ji} = \begin{cases} +1, & \text{with probability } \frac{1}{2}p \text{ or} \\ -1, & \text{with probability } \frac{1}{2}p \text{ or } \} \\ 0 & \text{otherwise} \end{cases} \tag{6.5.8}$$

and $G = \{G_a\}$ is used as the new variable of integration. Each such Γ^a corresponds

to some random multi-element excitation of the condensate. With the extreme choice $p = 0$ only one trivial Γ^a is formed. On the other hand, if $p = 1$ the Γ^a are formed with all off-diagonal entries being $+1$ or -1. However if $p \ll 1$ then the Γ^a have sparse non-zero entries, and approximate a complete set.

It turns out that this choice of the Γ set describes multi-element deviations from the condensate and has intrinsic characteristics that prompt a natural geometric interpretation.

6.6 Random graphs

Among the Pythagoreans the *monad* was the first thing that came into existence. The concept was fundamental to Pythagorean cosmology. Within certain variations of Gnosticism, the monad was the higher being which created lesser gods, or elements. In the metaphysics of Gottfried Leibniz, the monad was an indivisible, impenetrable unit of substance viewed as the basic constituent element of physical reality. Leibniz's theory does not posit physical space; rather, physical objects are constructs of the 'collective experiences' of monads. Leibniz was a leading advocate of the *relational* mode of thinking, both in response to and in contrast with Newton's *absolute* space and time.

Finding no more appropriate terminology than this, the word 'monad' is adopted here with much the same sense as that given above but with some important qualifications:

- a monad is not a geometrical object; in particular, it is not a mathematical point – it is not a zero-dimensional object embedded in a given background

geometric structure

- monads are purely conceptual entities having no objective existence, they are not 'things'
- a monad is just a member of a countable and abstract set

There is no sense in which monads may be better defined at this stage. While it may be convenient to share Leibniz's notion of monads as the basic constituent of physical reality this, if taken literally, would be a fundamental misconception. Monads acquire their meaning not by objective definition but by their functional relationship in the larger scheme, as will be seen.

Suppose that there exists a very large (possibly infinite) number, M, of monads. Further suppose that each monad 'experiences', to varying degree, the presence or existence of every other monad and that these 'experiences' may be represented by pair-wise link variables, m_{ij}, of corresponding magnitude with the following properties:

- magnitudes follow a Gaussian distribution (following the example of SQM, which employs Gaussian white noise – see §2.3.2 on page 60. Gaussian white noise is also used to model thermal noise and is the simplest idealization of noise [258])
- antisymmetry, so that $m_{ij} = -m_{ji}$;
- self-interaction is excluded, so that $m_{ii} = 0$.

There is no background geometry, nor any topology. Nevertheless, a straightforward way to introduce and define a natural measure of 'distance' between points is

to establish an adjacency relationship between the monads, specified for all possible pairs because the absence of background or embedding geometry means that the adjacency (or otherwise) of a given pair of monads is independent of the adjacency of any other pair. This obtains by identifying the property of adjacency with the 'strongest' links – those with the greatest absolute value. Next, coarse-grain or partition the link variables such that monads in any given pair are adjacent with probability p and not adjacent with probability $q = 1 - p$, with the requirement that $p \ll 1$.

The described relationships are readily identified with the Γ^a matrices, to each of which may be associated a random graph: let the indices i and j in Γ^a_{ij} label the *vertices* (synonymously, 'points' or 'nodes') of a graph in which two vertices i and j are connected by an *edge* (synonymously, 'link' or 'bond') – i.e. they are *adjacent* – if $\left|\Gamma^a_{ij}\right| = 1$.

6.6.1 Graph Theoretic definitions

Since its inception, the language of graph theory has been clouded by uncertain or ambiguous terminology. The following definitions[2] are considered to be consistent with the terminology and conventions of modern graph theory and are presented here to aid clarity in the arguments that follow.

Definition 6.6.1. (order): *The number of vertices of a graph is its* order, *usually written as* $|G|$.

Remark 6.6.1.1. Graphs are *finite* or *infinite* according to their order.

[2]The definitions obtain, largely but not entirely, from the Wikipedia *Glossary of graph theory* on the World Wide Web at http://en.wikipedia.org/wiki/Glossary_of_graph_theory

Definition 6.6.2. (degree): *The* degree *or* valency $d(v)$ *of a vertex* v *is the number of edges at* v.

Definition 6.6.3. (adjacency): *Two vertices* u *and* v *are considered* adjacent *if an edge exists between them. Adjacency is denoted by* $u \downarrow v$.

Definition 6.6.4. (nearest neighbours): *Mutually adjacent nodes are known, synonymously, as* nearest neighbours.

Remark 6.6.4.1. According to the usage here (but not generally so) the number of nearest neighbours of v is equal to the degree of v.

Remark 6.6.4.2. The terms *next-nearest neighbour* and so forth carry the obvious and intuitive interpretation.

Definition 6.6.5. (neighbourhood): *The set of neighbors, called an* (open) neighborhood $N_G(v)$ *for a vertex* v *in a graph* G*, consists of all vertices adjacent to* v *but not including* v*. When* v *is also included, it is called a* closed neighborhood*, denoted by* $N_G[v]$*. When stated without any qualification, a neighborhood is assumed to be open.*

Definition 6.6.6. (adjacency matrix): *The* adjacency matrix *of a directed or undirected graph with* n *vertices is the* $n \times n$ *matrix whose entry in row* i *and column* j *gives the number of edges from the* i^{th} *to the* j^{th} *vertex.*

Remark 6.6.6.1. An *adjacency matrix* is known, synonymously, as a *connectivity matrix.*

Remark 6.6.6.2. Unless specified otherwise, *adjacency matrix* means the *modified adjacency matrix.*

Definition 6.6.7. (modified adjacency matrix): *The* modified adjacency matrix *is generated by replacing all entries greater than 1 in the adjacency matrix by 1.*

Definition 6.6.8. (connectivity): *If it is possible to establish a path from any vertex to any other vertex of a graph, the graph is said to be* connected*; otherwise, the graph is* disconnected. *A graph is* totally disconnected *if there is no path connecting any pair of vertices.*

Definition 6.6.9. (path): *a* path *in a graph is a sequence of vertices such that from each vertex in the sequence there is an edge to the successor vertex. (Equivalently, a path is a sequence of vertices such that each vertex in the sequence is adjacent to both its predecessor and successor vertices.)*

Definition 6.6.10. (simple path): *A path is called a* simple path *if none of the vertices in the path are repeated.*

Definition 6.6.11. (cycle): *a* cycle *of n vertices is assumed to be a simple cycle, or a simple circuit in the modern sense, meaning every vertex is incident to exactly two edges.*

Remark 6.6.11.1. Traditionally, a cycle of a graph consisted of a sequence of successively incident edges and their end-vertices, where the terminating vertices are identical. In modern literature, that definition usually refers to what is known as a *circuit*, or *closed walk.*

Definition 6.6.12. (unit distance): *If two nodes are nearest neighbours then they are separated by a* unit distance *of 1 unit of length.*

Definition 6.6.13. (distance): *The* distance $d_G(u, v)$ *between two (not necessarily distinct) vertices u and v in a graph G is the length of a shortest path between them.*

Remark 6.6.13.1. This distance has *metric* properties since

$$\begin{aligned} d(u,u) &= 0, \\ d(u,v) &= d(v,u) \\ d(u,w) &\leq d(u,v) + d(v,w) \end{aligned} \tag{6.6.1}$$

(the subscript G is usually dropped when there is no danger of confusion). When u and v are unreachable from each other, their distance is defined to be infinity ∞.

Definition 6.6.14. (subgraph): *A* subgraph *of a graph G is a graph whose vertex and edge sets are subsets of those of G.*

Definition 6.6.15. (forest): *A* forest *is an acyclic graph.*

Definition 6.6.16. (tree): *A* tree *is a graph in which any two vertices are connected by exactly one path. That is, a tree T is minimally connected, so that $T - e$ is disconnected for every edge e in T (i.e. $T - e$ is a forest).*

Remark 6.6.16.1. A connected graph with n vertices is a tree if and only if it has $n - 1$ edges.

Remark 6.6.16.2. The vertices of a tree can always be enumerated, say as $v_1, \ldots, v_n$ so that every v_i with $i \geq 2$, has a unique neighbour in $\{v_1, \ldots, v_{i-1}\}$.

Remark 6.6.16.3. Given N different vertices, there are N^{N-2} different ways to connect them to make a tree.

Definition 6.6.17. (root node): *A* root node r *of a tree graph T is a vertex chosen so as to be regarded as the origin of any partial ordering on the set of vertices $V(T)$.*

Remark 6.6.17.1. r is the least element in such a partial order.

Definition 6.6.18. (leaf node): *A* leaf node *is any vertex of degree 1.*

Remark 6.6.18.1. Every non-trivial tree has at least two leaves – for instance, the ends of a longest path. If a leaf is removed, what remains is still a tree whereas removing a vertex of degree $d(v) > 1$ results in a forest.

Remark 6.6.18.2. Where a root vertex r has been chosen, thus imposing a partial ordering, every leaf $x \neq r$ of T is a maximal element of that order.

Definition 6.6.19. (spanning tree): *A* spanning tree *of a graph G is a tree that contains every vertex of G and whose edges are edges of G.*

Definition 6.6.20. (minimum spanning tree): *A* minimum spanning tree *spans the graph and is a minimum – the total weight of all the edges is as low as possible.*

Remark 6.6.20.1. If all paths are weighted equally (e.g. as in HQS) then the distance between any two vertices is the shortest possible distance between those points.

Returning to the Γ^a matrices, it follows from the above that each is a modified adjacency matrix of some random graph, G^a. Moreover, in general each graph is disconnected, being partitioned into a number of mutually disconnected components, or subgraphs – that is, there is a path of finite length between any two given vertices in the same component, while no such path exists between vertices in different components.

For graphs such as these, with probability of adjacency $p \ll 1$, Nagels [352] has considered the probability distribution of the distances between the vertices of connected subgraphs, first observing that although the assignment of adjacencies takes place independently for each possible pair of monads, the distances between monads are no longer independent; for instance, if i is adjacent to both j and k then

$d(j,k) = 1$ with probability p if j and k are nearest neighbours otherwise $d(j,k) = 2$ since j and k must, in that case, be next-nearest neighbours with the shortest path between them obtaining via i.

6.6.2 Review of Nagels

Nagels' [352] findings are central to the present argument, thus it is both relevant and informative to review and validate the key results, beginning with Eq.(5) in [352]. First, while it is assumed that the total number of monads is very large, the number of monads in any connected subgraph is certainly finite. The entire set M is thus partitioned into finite subsets of mutually disconnected components, each having $N_i, i \to \infty$ monads, which are at least simply connected – that is, each N_i may be described by a non-directed random graph where the monads are represented as vertices and the adjacency relations between these vertices, on a pairwise basis, provide edges of the graph. The minimal such graphs must then have $N_i - 1$ edges i.e. they are tree-graphs. Select one of these components, with $N = N_i \gg 1$, and choose, *arbitrarily*, one vertex to be the origin - a point of reference for the whole graph – and examine how the distribution of vertices appears with respect to that chosen origin, *viz*:

Let D_k be the number of vertices at a distance k from the origin then, as figure 6.2 on the next page illustrates, $D_0 = 1$ is the origin, D_1 is the number of adjacent vertices or nearest neighbours to the origin, and D_2 is the number of next nearest neighbours and so on.

Since N is finite, there is a maximum distance L on the graph and D_L is the number of vertices at this maximum distance from the origin. Clearly, then, there is

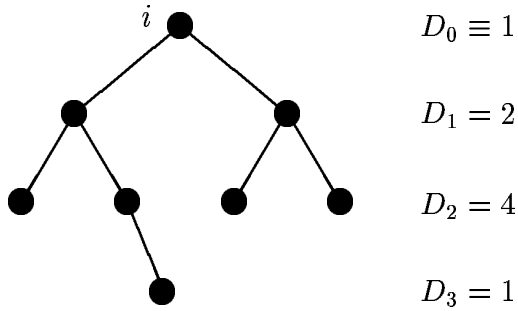


Figure 6.2: Example of a tree graph
($N = 8$, $L = 3$) with distance distributions from i^{th} node.

the constraint that

$$\sum_{k=0}^{L} D_k = N, \tag{6.6.2}$$

and also

$$\begin{cases} D_0 = 1 \\ D_k > 0, \quad 0 \leq k \leq L \\ D_k = 0, \quad k > L \end{cases} . \tag{6.6.3}$$

No vertex may be more than a distance $N - 1$ from the origin since $L = N - 1$ provides the simplest connected graph, a chain with $D_k = 1 \ \forall \ \{k : 0 \leq k \leq L\}$.

To calculate the probability $\{D_k : \ 0 \leq k \leq N, \ \sum_{k=0}^{N-1} D_k = N\}$:

1. the probability for the number D_1 of nearest neighbours (i.e. those vertices at unit distance from the origin) is p^{D_1}, which may be written as $(1-q)^{D_1} = (1 - q_0^{D_0})^{D_1}$, since $D_0 = 1$;

2. the probability for the next nearest neighbours, D_2, is obtained by considering that any vertex at this level is

   (a) adjacent to at least one point at unit distance from the origin;

(b) not adjacent to the origin itself.

Condition (b) is easily obtained since it occurs with probability $q = 1 - p$ so there is a factor of q^{D_2} for this.

Condition (a) may be obtained by first considering the counter argument i.e. that the vertex is *not* adjacent to any of the D_1. This has probability q^{D_1}. Thus the probability that it *is* adjacent to at least one of the D_1 is just $1 - q^{D_1}$. So there is an overall factor of $(1 - q^{D_1})^{D_2}$ for this condition.

Hence, the probability of obtaining D_2 is the product of these two factors

$$\text{prob}\,(D_2) = \left(1 - q^{D_1}\right)^{D_2}\ q^{D_2}; \tag{6.6.4}$$

3. the probability for D_3, those vertices at distance $k = 3$ from the origin, is similarly defined by the requirements that a vertex in D_3 is

   (a) adjacent to least one vertex in D_2;

   (b) not adjacent to any vertex in D_1;

   (c) not adjacent to the origin.

   Condition (a) is argued precisely as the corresponding condition in item 2 above, i.e. it provides a factor $(1 - q^{D_2})^{D_3}$.

   Condition (b) is expressed as q^{D_1}, thus providing the factor $(q^{D_1})^{D_3}$.

   Condition (c) is satisfied simply by the factor q^{D_3}, which may be written as $(q^{D_0})^{D_3}$ since $D_0 \equiv 1$. Hence the probability of obtaining D_3 is

$$\left(1 - q^{D_2}\right)^{D_3} \left(q^{D_1}\right)^{D_3} \left(q^{D_0}\right)^{D_3} = \left(q^{D_0+D_1}\right)^{D_3} \left(1 - q^{D_2}\right)^{D_3}; \tag{6.6.5}$$

4. for vertices at a distance $i+1$ from the origin, mathematical induction on the previous results gives

$$\text{prob}\,(D_{i+1}) = \left(q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}} \left(1 - q^{D_i}\right)^{D_{i+1}}. \tag{6.6.6}$$

So, the probability P on the whole graph (ie for a specified shape) is the probability of obtaining a particular set $(D_1,\ D_2,\ \cdots,\ D_L)$ which is

$$P = p^{D_1} \prod_{i=1}^{L-1} \left(q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}} \left(1 - q^{D_i}\right)^{D_{i+1}}. \tag{6.6.7}$$

Finally, note that vertices may be permuted between the *sets* of vertices at different distances. That is, the same magnitudes for each D_k could be obtained by many other possible configurations which result from a relabelling of the graph. First, there are $(N-1)!$ ways of relabelling the graph once the choice of origin has been fixed so there are $(N-1)!$ ways of obtaining the same P, where the depth structure given by $(D_1,\ D_2,\ \cdots,\ D_L)$ is identical. Second, the number of instances of a particular shape irrespective of labelling (beyond the choice of origin) is given by the product $D_1!D_2!\ \cdots\ D_L!$.

Hence there are $\frac{(N-1)!}{D_1!D_2!\ \cdots\ D_L!}$ ways of obtaining a graph (from a fixed origin) with a particular depth structure and therefore, the probability for a specified shape with N given and the origin arbitrarily chosen, that is, the probability distribution, is

$$\mathcal{P} = \frac{(N-1)!}{D_1!D_2!\ \cdots\ D_L!}\, p^{D_1} \prod_{i=1}^{L-1} \left(q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}} \left(1 - q^{D_i}\right)^{D_{i+1}}. \tag{6.6.8}$$

Now,

$$\begin{aligned} \left(q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}} \left(1 - q^{D_i}\right)^{D_{i+1}} &= \left(q^{\sum_{j=0}^{i-1} D_j} - q^{D_i} q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}} \\ &= \left(q^{\sum_{j=0}^{i-1} D_j} - q^{\sum_{j=0}^{i} D_j}\right)^{D_{i+1}}. \end{aligned} \tag{6.6.9}$$

So, finally, the expression for $\mathcal{P}$ is written as

$$\mathcal{P} = (N-1)!\,\frac{p^{D_1}}{D_1!}\,\prod_{i=2}^{L}\frac{1}{D_i!}\left(q^{\sum_{j=0}^{i-2}D_j} - q^{\sum_{j=0}^{i-1}D_j}\right)^{D_i}, \tag{6.6.10}$$

thus Nagels' equation (5) is derived and validated as required.

Expanding (6.6.10) in powers of p, the first two terms are

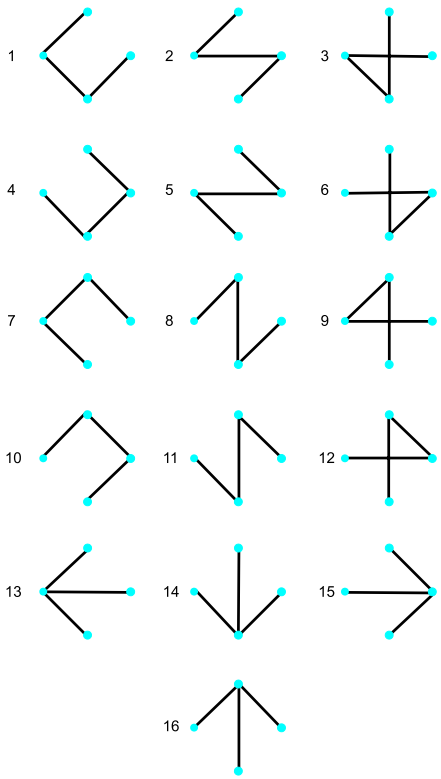
$$\mathcal{P} = \begin{cases} p^{N-1}, & L = 1 \\ p^{N-1}\,(N-1)!\left(\prod_{i=1}^{L}\frac{{D_{i-1}}^{D_i}}{D_i!}\right)\left\{1 + \frac{1}{2}p\left[\sum_{i=1}^{L}{D_i}^2\right.\right. & \\ \qquad \left.\left. +\sum_{i=1}^{L}D_iD_{i-1} - N^2 + N\right] + \mathrm{O}(p^2)\right\}, & L > 1 \end{cases} \tag{6.6.11}$$

and for fixed N with $p \ll 1$, the 'shape' of the graph (i.e. in terms of its depth structure, specified by the distribution of the D_is) may be determined by considering only the leading term of (6.6.11) so that

$$\mathcal{P}_1 = \prod_{i=1}^{L}\frac{{D_{i-1}}^{D_i}}{D_i!}. \tag{6.6.12}$$

Although it is convenient to continue the use of the label '$\mathcal{P}$' in the sense of a probability, the quantity $\mathcal{P}_1$ should not be mistaken for a probability measure in a literal sense – rather, it is a counting measure on a tree graph, being the relative frequency of a particular shape. The same result may be obtained more directly by considering the point of view of one monad, call it monad i, which is connected to a number of other monads with $p \ll 1$. The connections are represented graphically and because these links are very improbable the most likely form is that of a tree graph, which, though rare, is much more likely than a similar graph with cycles. (Moreover, even in the case where cycles are present, the depth structure of a graph is identical to the depth structure of a minimum spanning tree over that graph.)

The number of different possible linkage patterns between adjacent 'levels' k and $k+1$ is $D_k^{D_{k+1}}$ and so there are $D_1^{D_2} D_2^{D_3} \ldots D_{L-1}^{D_L}$ instances of a specific linkage pattern for the whole graph, which is not the same thing as the 'shape' of the graph – from the point of view of the root node, shapes are unique but linkage patterns are not – there are fewer shapes than linkage patterns, as the simple case shown in figure 6.3 demonstrates.



$(N-1)! = 6$

$L = 3$: $\{D_k\} = \{1,1,1\}$
(1 - 6) $\prod {D_{k-1}}^{D_k} = 1$
$\prod D_k! = 1$
$\mathcal{D}(D,N) = 6 \times 1/1 = 6$

$L = 2$: $\{D_k\} = \{1,2,1\}$
(7 - 12) $\prod {D_{k-1}}^{D_k} = 2$
$\prod D_k! = 2$
$\mathcal{D}(D,N) = 6 \times 2/2 = 6$

$L = 2$: $\{D_k\} = \{1,1,2\}$
(13 - 15) $\prod {D_{k-1}}^{D_k} = 1$
$\prod D_k! = 2$
$\mathcal{D}(D,N) = 6 \times 1/2 = 3$

$L = 1$: $\{D_k\} = \{1,3\}$
(16) $\prod {D_{k-1}}^{D_k} = 1$
$\prod D_k! = 6$
$\mathcal{D}(D,N) = 6 \times 1/6 = 1$

Figure 6.3: All possible tree-graph shapes for $N = 4$
showing the distinction between linkage pattern and the shape $\{D_k\}$. Counting the shapes proceeds according to equation 6.6.15.

Now consider the number of different N-node trees, with the same distance distribution $\{D_k\}$, to which i can belong. Given M monads from which to choose i, there are $\binom{M-1}{N-1}$ ways to obtain the other $(N-1)$ monads in the connected subgraph

that comprises the tree and $(N-1)!$ ways of permuting or re-labelling them, giving a factor[3] of

$$\begin{aligned}(N-1)!\,\tbinom{M-1}{N-1} &= \frac{(M-1)!\,(N-1)!}{(N-1)!\,(M-N-2)!} \\ &= \frac{(M-1)!}{(M-N-2)!}.\end{aligned} \tag{6.6.13}$$

Thus, counting different linkage patterns, together with permutations of the monads provides

$$\begin{aligned}\mathcal{N}(D_1,\ldots,D_L,N) &= \frac{(M-1)!\;D_1^{D_2}D_2^{D_3}\ldots D_{L-1}^{D_L}}{(M-N-2)!\;D_1!D_2!\ldots D_L!} \\ &= \frac{(M-1)!}{(M-N-2)!}\prod_{k=1}^{L}\frac{{D_{k-1}}^{D_k}}{D_k!}\end{aligned} \tag{6.6.14}$$

where the product term is identical to $\mathcal{P}_1$ in (6.6.12).

In fact, if the subgraph status is disregarded and the tree-graph considered in its own right, then

$$\mathcal{N}_1(D,N) = (N-1)!\prod_{k=1}^{L}\frac{{D_{k-1}}^{D_k}}{D_k!} \tag{6.6.15}$$

exactly counts the number of occurrences of any given shape $\{D_k\}$ in the set of N^{2N} possible tree graphs with N vertices.

The most probable 'shape' is determined by maximizing $\mathcal{P}$ by varying the D_is subject to the conditions imposed by (6.6.2) and (6.6.3).

It is convenient to work with $\ln\mathcal{P}$ instead of $\mathcal{P}$ itself. Using the Lagrange multiplier method, let

$$F = \ln\mathcal{P} + \mu\left(\sum_{i=0}^{L}D_i - N\right) \tag{6.6.16}$$

[3]In the case of a Grassmann algebra G_{2N} with $M = 2N$ elements, the pre-factor becomes $(2M-1)!/(2M-N)!$.

where $\mathcal{P}$ is given by 6.6.10). Setting $\partial F/\partial D_k = 0$, consider

$$\frac{\partial F}{\partial D_k} = \frac{\partial}{\partial D_k}(\ln \mathcal{P}) + \frac{\partial}{\partial D_k}\left(\mu\left(\sum_{i=0}^{L} D_i - N\right)\right), \tag{6.6.17}$$

and observe:

1. trivially, the second term on the RHS gives μ only;

2. expanding the first term on the RHS gives:

$$\begin{aligned}\frac{\partial}{\partial D_k}(\ln \mathcal{P}) = \frac{\partial}{\partial D_k}\Bigg\{&\ln(N-1)! - \ln(D_1!\, D_2! \cdots D_L!) + D_1 \ln p \\ &+ \ln\left[\prod_{i=1}^{L-1}(q^{\sum_{j=0}^{i-1} D_j})^{D_{i+1}}(1-q^{D_i})^{D_{i+1}}\right]\Bigg\}.\end{aligned} \tag{6.6.18}$$

Now, working through the terms on the right of the above equation...

(a)

$$\frac{\partial}{\partial D_k} \ln (N-1)\,! = 0\,;$$

(b)

$$\frac{\partial}{\partial D_k} \ln (D_1!\, D_2! \cdots D_L!).$$

Expanding, gives

$$\frac{\partial}{\partial D_k}\Big\{\ln(D_1!) + \ln(D_2!) + \cdots + \ln(D_k!) + \cdots + \ln(D_L!)\Big\}$$

of which, only the k^{th} term survives differentiation, *viz*:

$$\frac{\partial}{\partial D_k} \ln (D_k!),$$

which is the digamma function, $\psi(z+1) = \Gamma'(z+1)\,/\,\Gamma(z+1)$ (so allowing D_k to take non-integer values, which is an acceptable approximation if D_k

is large) thus this term may be written simply as

$$\frac{\partial}{\partial D_k}\left[-\ln\left(D_1!\,D_2!\,\cdots\,D_L!\right)\right] = -\,\psi(D_k+1)\,;$$

(c)

$$\frac{\partial}{\partial D_k}\left(D_1 \ln p\right) = \begin{cases} 0, & \text{if } k \neq 1 \\ \ln p, & \text{if } k = 1 \end{cases};$$

(d)

$$\ln\left\{\prod_{i=1}^{L-1}\left(q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}}\left(1-q^{D_i}\right)^{D_{i+1}}\right\}$$

$$= \ln\left(\prod_{i=1}^{L-1}\left(q^{\sum_{j=0}^{i-1} D_j}\right)^{D_{i+1}}\right) + \ln\left(\prod_{i=1}^{L-1}\left(1-q^{D_i}\right)^{D_{i+1}}\right).$$

i. Re-writing the first term gives the expression

$$\sum_{i=1}^{L-1} D_{i+1}\,\ln\left(q^{\sum_{j=0}^{i-1} D_j}\right),$$

which is to be differentiated with respect to D_k. Expanding:

$$\frac{\partial}{\partial D_k}\Big[D_2 \ln q^{D_0} + D_3 \ln q^{D_0+D_1} + \cdots + D_k \ln q^{D_0+\cdots+D_{k-2}}$$

$$+ D_{k+1} \ln q^{D_0+\cdots+D_{k-1}} + D_{k+2} \ln q^{D_0+\cdots+D_k} + D_{k+3} \ln q^{D_0+\cdots+D_{k+1}}$$

$$+ \cdots + D_L \ln q^{D_0+\cdots+D_k+\cdots+D_{L-2}}\Big].$$

For $k \neq L$, terms with coefficient D_i, $i < k$, all vanish under differentiation as does the term with coefficient D_{k+1}. This leaves

$$(D_0 + D_1 + \cdots + D_{k-2}) \ln q + (D_{k+2} + \cdots + D_L) \ln q$$

$$= \sum_{i=0}^{L} D_i \ln q - (D_{k-1} + D_k + D_{k+1}) \ln q, \ (k \neq L).$$

When $k = L$, differentiation does not yield the coefficient D_{k+1} seen in the above. However, by definition D_{L+1} is identically zero, so the term may be retained without loss of generality. Also, $\sum_{i=0}^{L} D_i = N$, since this is the constraint condition. Thus

$$\frac{\partial}{\partial D_k} \ln \left(\prod_{i=1}^{L-1} \left(q^{\sum_{j=0}^{i-1} D_j} \right)^{D_{i+1}} \right) = N \ln q - (D_{k-1} + D_k + D_{k+1}) \ln q.$$

ii. Similarly, re-writing the second term gives

$$\ln \left(\prod_{i=1}^{L-1} \left(1 - q^{D_i}\right)^{D_{i+1}} \right) = \sum_{i=1}^{L-1} D_{i+1} \ln \left(1 - q^{D_i}\right).$$

Again, this is to be differentiated with respect to D_k so, as before, expand thus:

$$\frac{\partial}{\partial D_k} \left[D_2 \ln \left(1 - q^{D_1}\right) + D_3 \ln \left(1 - q^{D_2}\right) + \cdots \right.$$
$$+ D_k \ln \left(1 - q^{D_{k-1}}\right) + D_{k+1} \ln \left(1 - q^{D_k}\right) + \cdots$$
$$\left. + D_L \ln \left(1 - q^{D_{L-1}}\right) \right].$$

Here, in general only two terms survive the differentiation and so

$$\ln \left(1 - q^{D_{k-1}}\right) + D_{k+1} \frac{\partial}{\partial D_k} \ln \left(1 - q^{D_k}\right)$$
$$= \ln \left(1 - q^{D_{k-1}}\right) - D_{k+1} \frac{q^{D_k} \ln q}{1 - q^{D_k}}.$$

When $k = 1$ the first term on the right does not present; however, recall that in that instance the $\ln p$ term survives (see 2c). Since $p = 1 - q$ and by definition $D_0 = 1$, we can write $\ln p = \ln \left(1 - q^{D_0}\right) = \ln \left(1 - q^{D_{k-1}}\right)$, which compensates for the otherwise excluded term. At the other extreme, when $k = L$, the second term on the right vanishes ($D_{L+1} \equiv 0$.) Again, however, for clarity in the final expression the term may be retained without loss of generality.

The components of equation 6.6.17 are reassembled to show that

$$\frac{\partial F}{\partial D_k} = \ln\left(1 - q^{D_{k-1}}\right) - \psi(D_k + 1) - (D_{k-1} + D_k + D_{k+1}) \ln q$$
$$- D_{k+1}\frac{q^{D_k} \ln q}{1 - q^{D_k}} + N \ln q + \mu \quad = \quad 0,$$

which is further simplified to

$$\frac{\partial F}{\partial D_k} = \ln\left(1 - q^{D_{k-1}}\right) - \psi(D_k + 1) - (D_{k-1} + D_k) \ln q$$
$$- \left(D_{k+1} + D_{k+1}\frac{q^{D_k}}{1 - q^{D_k}}\right) \ln q + N \ln q + \mu.$$

The fourth term on the right is factored and simplified, i.e.

$$\left(D_{k+1} + D_{k+1}\frac{q^{D_k}}{1 - q^{D_k}}\right) \ln q \quad = \quad D_{k+1}\left(1 + \frac{q^{D_k}}{1 - q^{D_k}}\right) \ln q$$
$$= \quad D_{k+1}\frac{\ln q}{1 - q^{D_k}},$$

and thus finally to arrive at

$$\frac{\partial F}{\partial D_k} = \ln\left(1 - q^{D_{k-1}}\right) - \psi(D_k + 1) - (D_{k-1} + D_k) \ln q$$
$$- D_{k+1}\frac{\ln q}{1 - q^{D_k}} + N \ln q + \mu \quad = \quad 0, \tag{6.6.19}$$

which is Nagels' equation 9 from [352], now verified.

Continuing the assumptions $p \ll 1$, $D_k \gg 1$, with the further stipulation that the product pD_k may be neglected without loss of generality, (6.6.19) may be expanded and approximated to give the most probable D_k distribution in the recursion relation

$$D_{k+1} = D_k \ln\left(\frac{D_k}{D_{k-1}}\right) + D_k(1 - \eta). \tag{6.6.20}$$

For comparison, consider maximizing 6.6.15 in like manner. Again, using the convenience of the logarithmic form and employing the Lagrange multiplier method,

let

$$F' = \ln \mathcal{N}_1 + \mu \left(\sum_{i=0}^{L} D_i - N \right) \tag{6.6.21}$$

and let $\partial F'/\partial D_k = 0$. Now,

$$\frac{\partial F'}{\partial D_k} = \frac{\partial}{\partial D_k} (\ln \mathcal{N}_1) + \frac{\partial}{\partial D_k} \left(\mu \left(\sum_{i=0}^{L} D_i - N \right) \right). \tag{6.6.22}$$

As before, the second term on the RHS yields μ only. Expanding the first term gives:

$$\begin{aligned} \frac{\partial}{\partial D_k} (\ln \mathcal{N}_1) &= \frac{\partial}{\partial D_k} \ln [(N-1)!] + \frac{\partial}{\partial D_k} \sum_{i=1}^{L} (D_i \ln D_{i-1} - \ln(D_i!)) \\ &= \frac{\partial}{\partial D_k} (D_k \ln D_{k-1} + D_{k+1} \ln D_k - \ln(D_k!)) \\ &= \ln D_{k-1} + \frac{D_{k+1}}{D_k} - \frac{\partial}{\partial D_k} \ln(D_k!) \end{aligned} \tag{6.6.23}$$

Using Stirling's Approximation, $n! \approx \sqrt{2\pi n}(n/e)^n$ for sufficiently large n, the third term on the RHS provides

$$\begin{aligned} \frac{\partial}{\partial D_k} \ln(D_k!) &= \frac{\partial}{\partial D_k} \ln(\sqrt{2\pi}) + \frac{1}{2} \frac{\partial}{\partial D_k} \ln D_k + \frac{\partial}{\partial D_k} (D_k \ln D_k) - \frac{\partial}{\partial D_k} D_k \\ &= \frac{1}{2D_k} + \ln D_k + 1 - 1 \\ &= \ln D_k + \frac{1}{2D_k} \end{aligned} \tag{6.6.24}$$

Thus 6.6.22 becomes

$$\begin{aligned} \frac{\partial F'}{\partial D_k} &= \ln D_{k-1} + \frac{D_{k+1}}{D_k} - \ln D_k - \frac{1}{2D_k} + \mu = 0 \\ \Rightarrow \quad \frac{D_{k+1}}{D_k} - \frac{1}{2D_k} &= \ln \left(\frac{D_k}{D_{k-1}} \right) - \mu \end{aligned} \tag{6.6.25}$$

so arriving at the recursion relation (note the similarity with equation 6.6.20):

$$D_{k+1} = D_k \ln \left(\frac{D_k}{D_{k-1}} \right) - \mu D_k + \frac{1}{2} \tag{6.6.26}$$

Using, instead, that $\partial/\partial D_k \ln(D_k!)$ is the digamma function, would have led to the approximation $\ln D_k - \frac{1}{2D_k}$, with negligible effect on the result – the difference being only the sign on the last term above.

Neither 6.6.20 nor 6.6.26 yield a direct analytic solution but Nagels used perturbative methods for 6.6.20 to find the approximate solution (equation 30 in [352]):

$$D_k \simeq \frac{2N}{L}\left[\sin^2\left(\frac{\pi k}{L}\right) - \frac{1}{3}\left(\frac{\pi}{L}\right)\sin\left(\frac{2\pi k}{L}\right)\cdot \ln\sin\left(\frac{2\pi k}{L}\right)\right]. \tag{6.6.27}$$

More particularly, finding N in terms of L gives ([352], equation 43)

$$D_k \approx \frac{L^2 \ln L}{2\pi^2}\left[\sin^2\left(\frac{\pi k}{L}\right) - \frac{1}{3}\left(\frac{\pi}{L}\right)\sin\left(\frac{2\pi k}{L}\right)\cdot \ln\sin\left(\frac{2\pi k}{L}\right)\right]. \tag{6.6.28}$$

Here, the leading term dominates (unless k/L is smaller than $\ln(L)/L$) so that

$$D_k \propto \sin^2\left(\frac{\pi k}{L}\right). \tag{6.6.29}$$

with constant of proportionality such that

$$N \approx L^3\left(\frac{\ln L}{2\pi^2}\right). \tag{6.6.30}$$

6.7 From monads to *gebits*

The remarkable property revealed by the analysis of the most probable distribution in (6.6.28) is the appearance of the $\sin^2$ term, indicating that the tree-graph is embeddable in a 3-dimensional hypersphere, S^3. Most importantly, monad i 'sees' its surroundings as being 3-dimensional, since $D_k \sim k^2$ for small $\pi k/L$. Taken together with (6.6.30), this shows that the resulting emergent structure closely resembles a three-dimensional closed space of positive curvature.

Each random subgraph corresponds to a partitioning of Γ, so that for the integration variables G_a each subscript labels, in the most probable case, a relational structure specified by a Γ^a possessing intrinsic geometrical properties that characterize a three-dimensional closed space. If the exponent in (6.5.7) is expanded in powers of G_a, computing the trace summations leaves an induced effective action for the G_a

$$< \mathcal{N} \mid \mathcal{N} >= \int \mathcal{D}G e^{-S[G]} \tag{6.7.1}$$

corresponding to a quantum 'field' theory of interacting 3-spaces. While this is analogous to some programs that are being currently pursued in quantizing general relativity, it would be nâive to presume that these 3-spaces might be identified directly with the spatial section of a 'universe' – of themselves, the Γ^a structures lack sufficient richness, moreover, they are disjoint. In addition, the most likely distance distribution $\{D_k\}$, though possessing the $\sin^2$ signature, gives the subgraphs only weak embeddability in S^3. Strong embeddability, attainable only with the existence of extra cross-linking connections (equivalently, the presence of cycles in the random graphs), is required for the emergence of a 'true' 3-space and one might expect to see some further condensation processes for these 'core' 3-space states leading to a fractal or foamy spatial structure. Yet these rudimentary structures are at least indicative of an *incipient* 3-geometry corresponding to our customary notions of physical space. For this reason, these 3-spaces are named *gebits* (geometrical bits) [342].

6.8 Numerical studies of small-p random graphs

The foregoing analytic results, while robust, were subject to further scrutiny via various numerical investigations, which employed the *Mathematica* technical computing

software.

6.8.1 Relative frequency distribution of shape lengths L

For a random tree with N vertices, there are N^{N-2} possible graphs partitioned into 2^{N-2} 'shapes'. In principle, it is straightforward to explicitly list each possible distribution of nodes over all possible lengths of the graph (which range from 1 to $N-1$) – the list is simply an enumeration of the permutations of unrestricted partitions of the integer N. In practice, for even small N the task soon becomes prohibitive to carry out explicitly. Figure 6.4 on the current page illustrates the results for $N = 24$, which involved listing and inspecting 4,194,304 $\{D_k\}$ distributions. As may be expected, the distribution is highly symmetric and favours the median and near-median lengths.

In fact, the 'difficulty' of the task is illusory – for N even, these L-values are simply the binomial probability density function for $N - 2$.

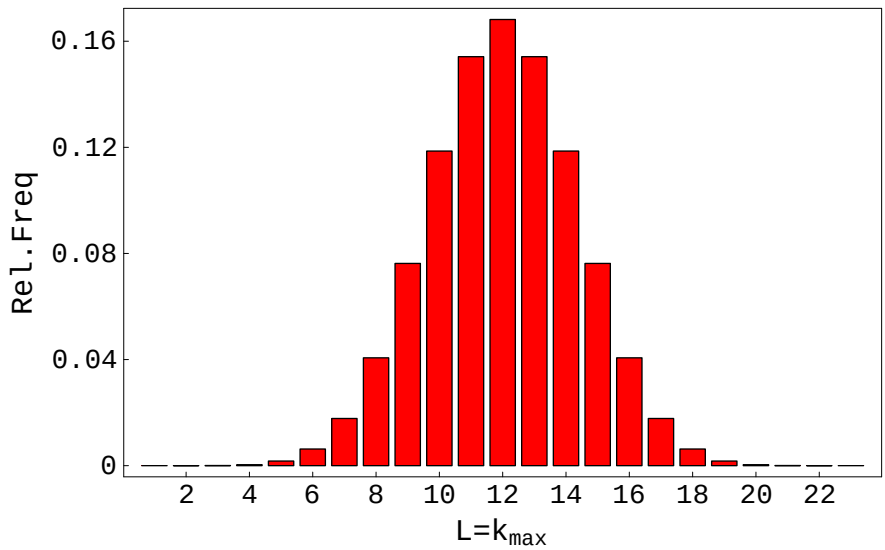


Figure 6.4: Probability distribution for graph length L over all possible $\{D_k\}$ shapes with $N = 24$

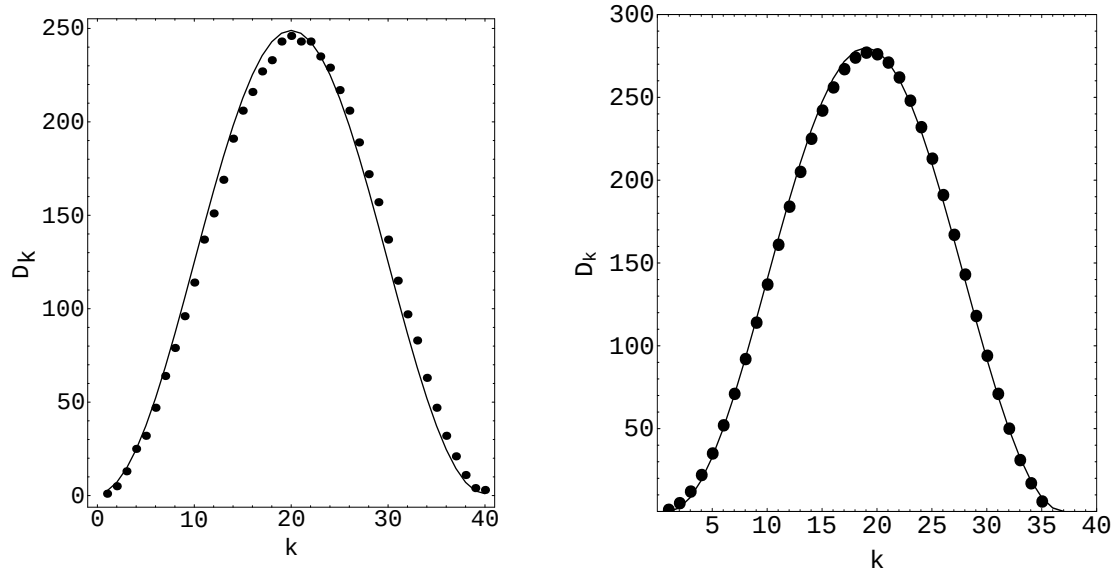
6.8.2 Numerical maximization of the pdf

To verify the analytic solution and find the most likely connection pattern, the probability distribution function (pdf) $\mathcal{P}[D, L, N]$ from equation 6.6.10 was numerically maximized with respect to L and $\{D_k\}$ for fixed $N = 5000$ and p such that $\log_{10}(p) = -6$. The resulting $\{D_k\}$ are plotted against the depth k in figure 6.5(a) on page 187, which shows the data are fitted very closely by the approximate analytic form $D_k \propto sin^2(\pi k/L)$ with $L = k_{\max}$. Similarly, $\mathcal{N}_1(D, N)$ – the tree-graph result of equation 6.6.15 – was numerically maximized, with the consequent $\{D_k\}$ likewise displayed in figure 6.5(b). These numeric results confirm the analytic findings that the most likely sub-graph to which a monad can belong has a distance distribution $\{D_k\}$ that is characteristic of the hypersphere S^3.

6.8.3 Gebit dimensionality as a function of link probability p

Figure 6.6 on page 188 shows the relationship between the intrinsic gebit dimensionality, d, and the link probability for a range of $\log_{10} p$ with $N = 5000$. In particular, note that $d = 3$ for all 'small' p but as p rises to approach and then exceed N^{-1}, so d increases sharply, making the graphs no longer embeddable in S^3. This threshold effect is an excellent illustration of the graph theoretic results that are described later in §6.9 on page 201.

The goodness of fit to the $\sin^2$ form demonstrated by figure 6.5 establishes the relationship between $\{D_k\}$, N and L inherent in the probability distribution function $\mathcal{P}$ and echoed in the simpler form of $\mathcal{N}_1$ for tree-graphs. Since this derived primarily from the requirement that $p \ll 1$ and held the consequence of the trigonometric form for the solution of the optimal D_k values, it prompts an investigation of the



(a) Case: $\mathcal{P}[D, L, N]$

(b) Case: $\mathcal{N}_1(D, N)$

Figure 6.5: Maximized shape distributions
Points show $\{D_k\}$ values against depth k. The curve $f(k, L) \propto sin^2(\pi k/L)$ is fitted to the plotted points in each case. 6.5(a) is the more general case from equation 6.6.10 with $N = 5000$ and $\log_{10}(p) = -6$, while 6.5(b) is the tree-graph result of equation 6.6.15 with $N \approx 5000$ and $\mu = 0.015$.

consequences to the HQS system of changes in p, in a sense coming full circle.

6.8.4 Examination of $\mathcal{P}_{\max}$ w.r.t. the $\sin^{d-1}$ form of $\{D_k\}$

Equation 6.6.10 provides the means for computing the (relative) probability for a specified shape $\{D_k\}$, while equation 6.6.28 shows that the most likely class of graph has the form $\sin^{d-1}$, with $d = 3$ representing the dimension of the space described by that form. However, the behaviour of (6.6.10) is far from obvious. Solving (6.6.30) for L gives

$$L \sim \sqrt[3]{\frac{12N\pi^2}{\ln N}} \tag{6.8.1}$$

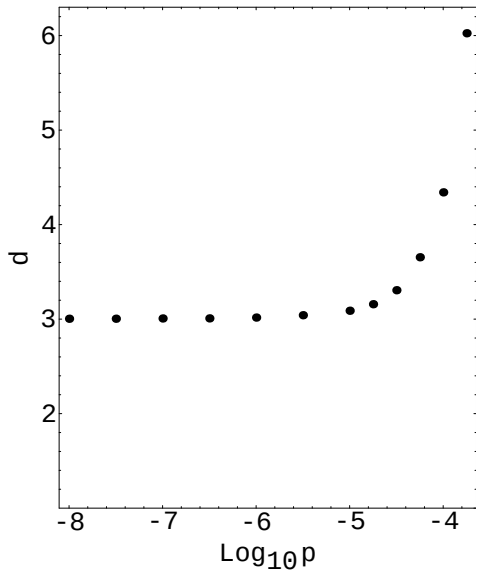


Figure 6.6: Gebit dimension versus link probability
Dimensionality d of gebits as a function of the link probability p, demonstrating a threshold effect (see §6.9 on page 201).

so that the choice of $N = 1000$ yields $L \approx 26$ as the length of the most likely shape. Taking $\{d : 2 \leq d \leq 6\}$ and $\{L : 15 \leq L \leq 35\}$, successive sets of $\{D_k\}$ were computed explicitly using a variation of (6.6.28), i.e.

$$D_k = \frac{N\sqrt{\pi}\,\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)L}\left[\sin^{d-1}\left(\frac{\pi k}{L}\right) - \text{Min}\left\{\frac{1}{3}\left(\frac{\pi}{L}\right)\sin\left(\frac{2\pi k}{L}\right)\cdot\ln\sin\left(\frac{2\pi k}{L}\right), 0\right\}\right]$$

and the results passed to (6.6.10) to find the probability of each D_k. The use of the Euler gamma function, $\Gamma(z)$ is necessary to accommodate the consequences of variations in d, keeping the sum of the D_ks sufficiently close to N. The results are shown in figure 6.7 on the next page, illustrating very clearly the dominance of $d \sim 3$ as the dimensional measure of the HQS gebits, as well as the equally strong indicator for the depth structure, or shape, evidenced by the L-dependency. Notice, too, that the maximum in L is in very good agreement with the theoretical value for the length of the most likely shape.

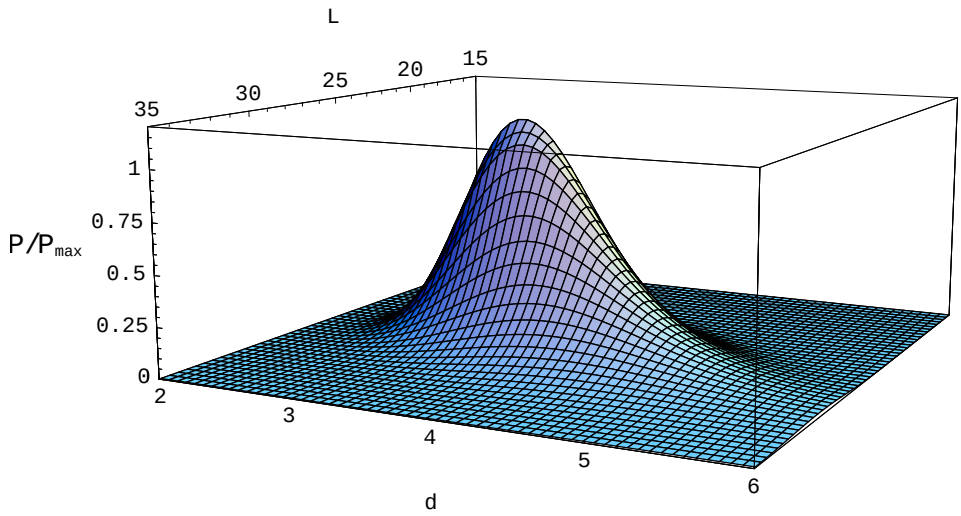


Figure 6.7: Probability distribution for $\{D_k\}$ as a function of d and L.

A particularly interesting feature, however, is the 'fuzziness' in the dimension – while the graph peaks around $d = 3$, clearly the favoured geometry is not single-valued, strongly suggesting that while the gebits have embeddability in a 3-dimensional hypersphere, S^3, they are not necessarily rigidly constrained to that form.

6.8.5 'Embeddability'

The issue of intrinsic dimensionality and what has been referred to as 'embeddability' is a fundamental issue for HQS theory and of vital importance to the increasingly sophisticated developments yet to be discussed. In general topology, an embedding is a homeomorphism of a space onto its image. That is to say, it can be thought of as an inclusion map on a space X that is considered as a subspace of Y, such that the topology on X is the same as the subspace topology it has in Y. Strictly, this is an extraneous notion in the current argument since it is only the emergence

of intrinsic dimensionality that matters here. For instance, a physical collection of linked point-like objects, the representation of which is a tree-graph, is *planar* because the structure is a subspace of the Euclidean 2-space. However, this is an extrinsic measure; intrinsically, the graph has its own measure, by virtue of the graph theoretic definitions given in §6.6.1, according to which the structure might be anything but planar. Nonetheless, the concept of embedding is a convenient and useful means of discussing and investigating intrinsic dimensionality since the identification of a homeomorphism appeals to intuition: the robustness of an emergent form may be gauged by the extent to which one is able to map points in the first object that are 'close together' to points in the second object that are close together and, similarly, points in the first object that are not close together to points in the second object that are not close together.

Trees and 'augmented' trees – weak *versus* strong embedding

The results thus far demonstrate the $\sin^2$ form as the graph shape with the highest probability, signifying S^3 as the most likely geometry 'seen' by an arbitrarily chosen monad. However, this view corresponds only to weak embeddability for gebits generally, for two principle reasons.

Firstly, while it is true that the $\sin^2$ form has the highest probability of any *particular* form, it is nonetheless a minority case – there are $(2^{N-2}-1)$ other available shapes which, taken collectively, provide many more opportunities for the occurrence of a graph shape that does *not* have the $\sin^2$ form.

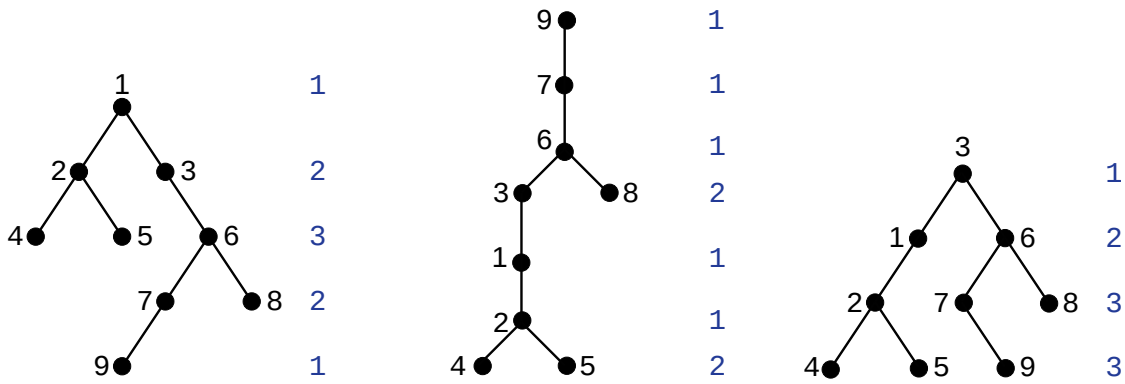
Secondly, the close agreement between the results of maximizing $\mathcal{P}[D, L, N]$ and $\mathcal{N}_1[D, N]$ affirms the dominance of tree sub-graphs over sub-graphs with cycles and

these must necessarily be only weakly embeddable since *strong* embeddability requires that every (or almost every) node in a graph, when taken as the root node, should yield a $\{D_k\}$ distribution consistent with that obtained from its neighbours when they, in turn, are considered as the root – that is, the graph should be highly symmetric. Consider, for instance, the symmetry of graphs representing the Platonic solids, regular lattices on open spaces, or uniform tessellations. A *uniform tessellation* is a tiling of a d-dimensional space, or a (hyper)surface, such that all its vertices are identical, i.e., there is the same combination and arrangement of faces at each vertex. Mathematically, a tiling of a topological space S consists of a collection B of open subsets of S, such that

- the shapes in B do not overlap (i.e., are mutually disjoint, having no point in common)
- they 'cover' S (the closure of their union is equal to S).

Each node of such graphs is of the same degree and is indistinguishable from every other node in the same graph. Clearly, tree graphs may vary widely in the symmetries they possess; just as clearly, they can never have the very high symmetry exhibited by Platonic solids and uniform tessellations. Consequently, inspecting the depth structure (that is, the nearest-neighbour counts and distances)[4] of an arbitrary tree-graph from the point of view of each node must yield multiple $\{D_k\}$ distributions, any one of which might legitimately be described as the shape of the graph. Figure 6.8 illustrates the situation in the simple case with $N = 9$. Each of the figures describes the same graph yet the depth structure is markedly different according to the choice of the root node.

[4] The depth structure is obtained through inspection of successive matrix powers of the graph's



(a) The most probable configuration

(b) The configuration according to node 9

(c) Likewise with node 3 as the root node

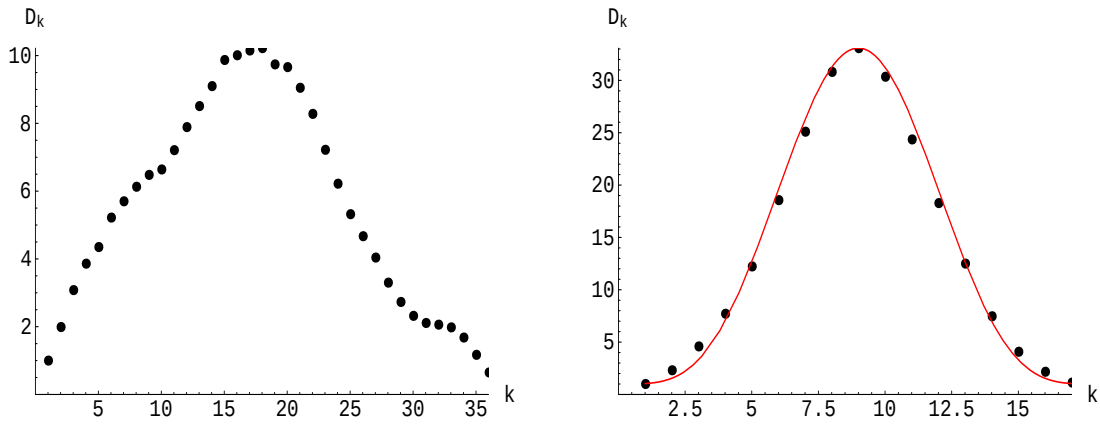
Figure 6.8: Root-node dependent tree shapes
Nodes are enumerated and D_k values are shown on the right of each example.

While the disparity of these node-dependent views is quite extreme for small graphs, numerical studies of larger graphs indicate that this problem tends to diminish somewhat with larger N due to the effect of the binomial probability density function on the graph length demonstrated in §6.8.1 on page 185. Since the median length dominates, there are vastly more shapes that are, on average, closer in form to the $\sin^2$ distribution. Although very many nodes in even a large graph will fail in this measure, averaging all N of the $\{D_k\}$ sets provides consistent results that tend (albeit, loosely) towards the desired form. Figure 6.9 on the next page shows a typical case for a single small random graph with $N = 200$ nodes. However, 6.9(a) demonstrates that

adjacency matrix, *viz:* if $A = \{a_{ij}\}_{n\times n}$ is the adjacency matrix of a graph G, vertex i in G has $d_{k,i}$ k^{th}-nearest neighbours where $d_{k,i}$ is the sum of unit-valued elements in the i^{th} row (excluding the i^{th} element) of the k^{th} matrix power of A, i.e.

$$d_{k,i} = \sum_{\substack{j=1 \\ j\neq i}}^{n} x_{ij}, \quad x_{ij} = \begin{cases} 1 & \text{if } \left(A^k\right)_{ij} = 1 \\ 0 & \text{otherwise} \end{cases}$$

for pure tree graphs, the disparate views continue to be problematic. The asymmetry in the plot arises because all tree-graph shapes must have at least nearest neighbours to the root node (so contributing to D_k for smaller k-values) whereas only the longest shapes provide D_k contributions for the larger k-values.



(a) Case: tree graph

(b) Case: augmented tree structure with $p = N^{-1.2} \approx 0.0017$

Figure 6.9: Average $\{D_k\}$ taken over all nodes of a single graph $N = 200$. The theoretical graph length is $L = 16$, using (6.8.1). 6.9(a) shows the original tree graph, containing no cycles. In 6.9(b), augmentation introduced 32 extra edges. The fitted curve is $f(k, L) \propto \sin^{\alpha}(\pi k/L)$ (in red).

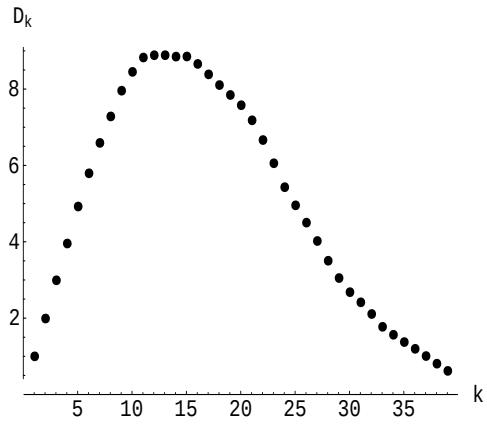
Fortunately, the more general case afforded by $\mathcal{P}[D, L, N]$ provides for the existence of cycles in the graph as p rises above the minimal level required to generate large tree graphs. A complete graph with n vertices has $\frac{1}{2}n(n-1)$ edges. A tree graph with n vertices has $(n-1)$ edges. The difference between these is $\frac{1}{2}n(n-1) - (n-1)$ $= \frac{1}{2}(n^2 - 3n + 2) \approx \frac{1}{2}n^2$ for large n. The probability of having an edge between any two vertices is $p \ll 1$ so for a random connected graph with n vertices (and at least $n-1$ edges) one would expect to find some $\frac{1}{2}pn^2$ 'extra' edges (i.e. the addition of

which generates cycles in the graph).

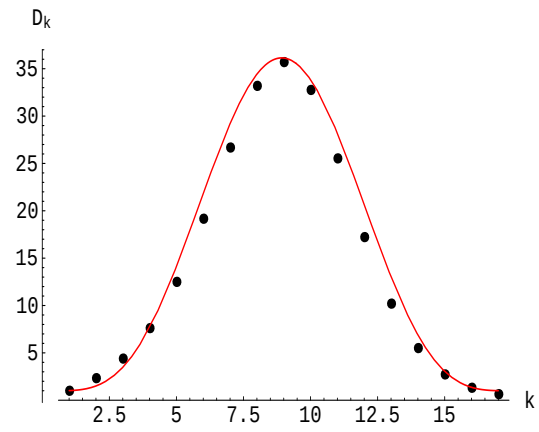
Figure 6.9(b) on the preceding page shows the consequence of augmenting the original tree by introducing additional random links. The regular fit of the curve $f(k, L) \propto \sin^{\alpha}(\pi k/L)$ (in red) clearly demonstrates the isotropy in the $\{D_k\}$ distribution that results from the introduction of cycles in the graph. In this particular case, the parameter $\alpha = 3.8$ is considerably higher than the desired value of 2 for strong embeddability in S^3. Moreover, while the graph length ($L = k_{\max}$) is in good agreement with theory ($L = 16$ for $N = 200$, using equation 6.8.1), the maximum D_k is too high – setting $k = L/2$ in equation 6.6.27 gives $D_{\max} = 2N/L = 25$ (using instead 6.6.28 would give $D_{\max} = L^2 \ln L/2\pi^2 = 36$ but this latter form is derived in the limit $N \to \infty$ and thus is considered unsuitable for small N). However, these factors will be seen to be artefacts of the simple averaging technique used, the choice of p-value for the augmentation procedure, and that the data obtain from a single sample.

The regularizing effect of the relatively sparse augmentation is not accidental. The procedure was repeated over a sample of one hundred random graphs with the same parameters as before. The very consistent results are shown in figure 6.10, where the parameter α has dropped to $\alpha = 3.4$, indicating that the higher value is indeed artefactual, although the peak D_k remains high.

To obtain the 'collective view' from many nodes and over many graphs as shown in figures 6.9 and 6.10, the respective $\{D_k\}$s were obtained by simple averaging – inspecting each graph from each node in turn taken as the root node to obtain the depth structure according to the method described in the footnote on page 191. The D_k values were then summed for each k and the results divided by the sample size.



(a) Case: tree graphs



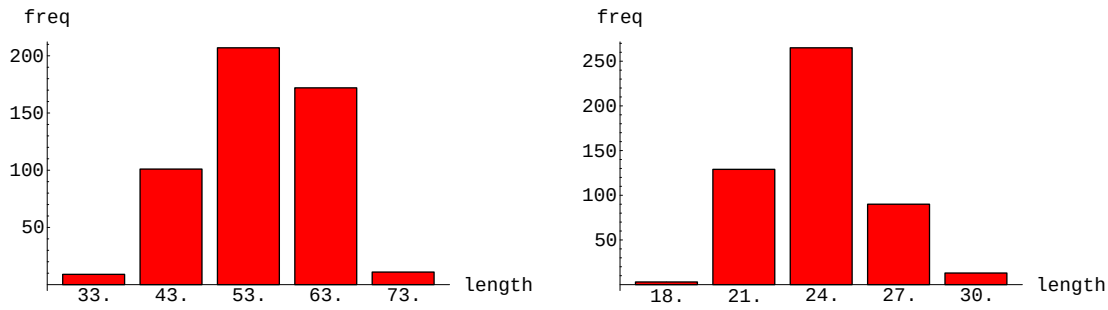
(b) Case: augmented tree structures

Figure 6.10: Simple average $\{D_k\}$ taken over all nodes of 100 graphs
As in figure 6.9, $N = 200$, $p = N^{-1.2}$ and the fitted curve in (b) is $f(k, L)$ with α here having the value 3.4

The ensuing 'averaged' $\{D_k\}$ is, however, distorted by the variation in graph lengths that arise as a consequence of the different views. A more sophisticated approach is to normalize both the vertex counts and the graph lengths and construct a histogram by sorting the data into bins. Finally, the data are re-scaled according to the average graph length and average maxima.

Statistic	Trees	Augmented trees
Mean	55.222	23.858
Median	55	24
Mode	58	23
Standard deviation	8.10754	2.10492
Mean deviation	6.59224	1.66252

Table 6.1: Descriptive statistics of graph length distributions
for $N = 500$, with $p = N^{-1.4}$ for the augmentation. See also figure 6.11.



(a) Case: trees

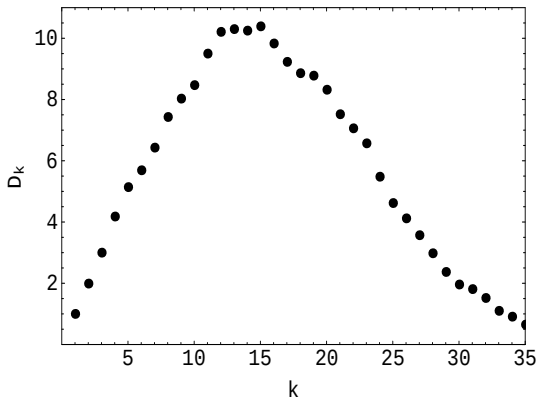
(b) Case: augmented trees

Figure 6.11: Histogram: distribution of graph lengths
for $N = 500$, $p = N^{-1.4}$ for the augmentation.

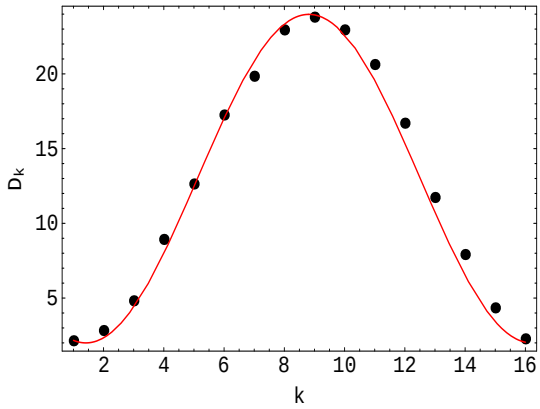
Table 6.1 on the preceding page contains descriptive statistic measures of the graph length distributions and these are further illustrated by the histograms in figure 6.11, which report the distribution of graph lengths for the simple tree graph case and then the augmented trees (respectively, 6.11(a) and 6.11(b)). The greater variability in the tree graph structures is evident from the range of lengths whereas the augmented tree lengths have relatively little spread and are grouped close to the theoretical value.

The results appear in figure 6.12 on the next page for a trial over 200 samples with $N = 200$ (6.12(a) and 6.12(b)) and $N = 500$ (6.12(c) and 6.12(d)). Also, the value p for the probability of extra links was optimized with respect to both the respective averaged graph lengths and maximum D_k values.

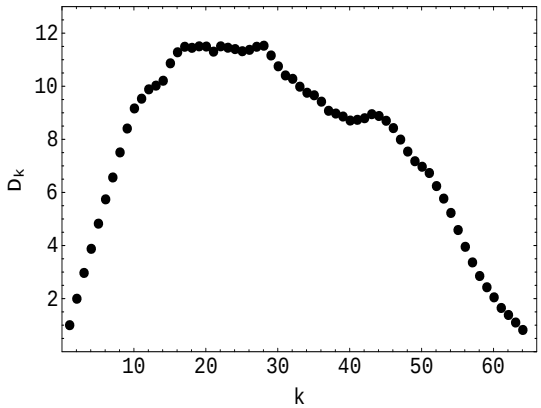
Notice there that both L and the peak D_k are in excellent agreement with the theoretical values (respectively, $L = 16$, $D_{max} = 24$ for $N = 200$ and $L = 21$, $D_{\max} = 47$ for $N = 500$) and, most notably, the fitted curves now have $\alpha = 2$. The greatest significance of these results, however, is the extent to which augmentation



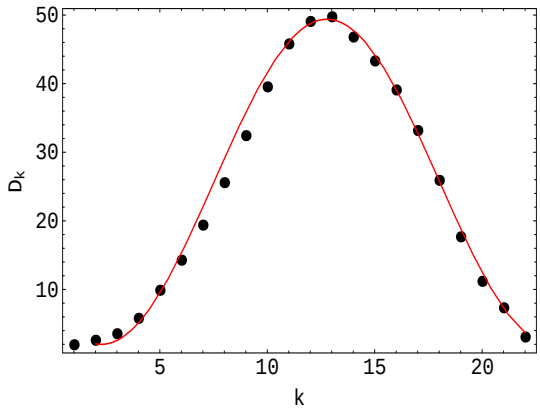
(a) Case: tree graphs



(b) Case: augmented tree structures

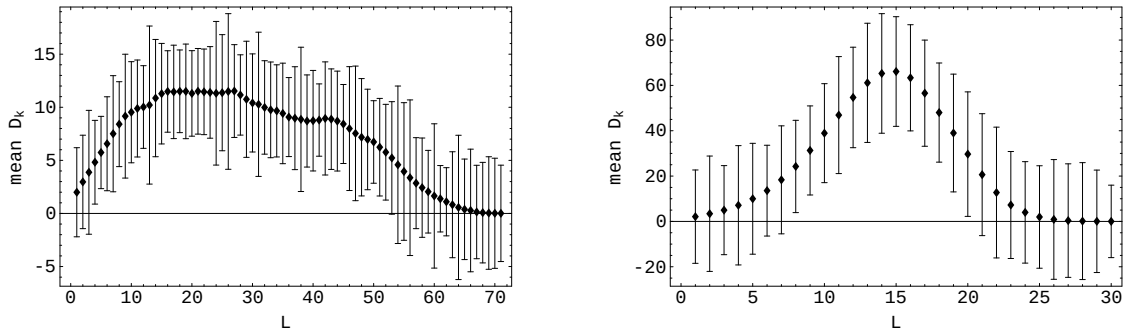


(c) Case: tree graphs



(d) Case: augmented tree structures

Figure 6.12: Average $\{D_k\}$ taken over all nodes of 200 graphs
(a) and (b) have $N = 200$ as before while (c) and (d) have $N = 500$. The augmentation was done with $p = N^{-1.4}$ and the fitted curves in (b) and (d) are $f(k, L)$ now with $\alpha = 2$ in each instance.



(a) Case: trees

(b) Case: augmented trees

Figure 6.13: Mean D_k values with error bars showing the $\pm$ mean (i.e. absolute) deviation. As before, $N = 500$, $p = N^{-1.4}$ for the augmentation.

leads all nodes to 'agree' on the intrinsic geometry, which turns out to be the S^3 form as predicted, demonstrating how the weak embeddability of tree graphs turns to strong embeddability with the appearance of a relatively small number of cycles. The sensitivity to the augmentation density via the p-value further suggests a *threshold effect* consistent with random graph theory (see §6.9 on page 201).

Figure 6.13 shows the mean D_k values obtained from the same sample with error bars indicating the mean deviation (otherwise known as the absolute deviation defined by $\frac{1}{n}\sum_{i=0}^{n-1}|x_i - \hat{x}|$) in each D_k. While the deviations are relatively large in both the tree and the augmented cases, they appear less variable in the latter. If *very* strong embeddability was to be a desirable property, one might be disappointed to not observe smaller variation; however, for present purposes the contrary observation is a nice result – gebits have strong but 'fuzzy' embeddability in S^3, which will be shown to have profound consequences.

6.9 Notes from random graph theory

If V is a fixed set of n elements, $V = \{0, \ldots, n-1\}$, a common aim in graph theory is to turn the set $\mathcal{G}$ of all graphs on V into a probability space $\mathcal{G}(n,p)$ and then question the probability that a graph $G \in \mathcal{G}$ has some particular property $\mathcal{P}$. Frequently, the value of p is related to the cardinality n of V, perhaps to the extent that $p = p(n)$ for a function $n \mapsto p(n)$ or merely the stipulation (such as in HPS, say) that $p \ll 1$ while $N \gg 1$. In standard probabilistic language, any set of graphs on V is called an *event* in $\mathcal{G}(n,p)$. The probability measure P is the product measure of all measures P_e, the probability of elementary events on every potential edge $e \in [V]^2$. P is uniquely determined by the value of p and by the provable assertion that events A_e are independent. In discrete mathematics, the existence of an object having a desired property may be proved by defining a probability space on a larger, non-empty class of objects and then showing the positive probability that an element of this space possesses the property desired. This is known as the *probabilistic method*, a classic example of which is Erdös's theorem on large girth and chromatic number. The 'objects' of the probability space may be partitions, mappings, embeddings – literally, objects of any kind. It is a well known result of random graph theory that all statements about graphs expressed by quantifying over vertices only, as distinct from sequences or sets of vertices, are consequences of asserting the existence of some graph property $\mathcal{P}_{i,j}$ for suitable i, j and that any such statement is almost certainly true or almost certainly false. Moreover, counter-intuitively, most such results share the common feature that the outcome is independent of the value of p – that is, if almost every graph in $\mathcal{G}(n, 0.5)$ possesses the property then the same tends to be true for almost every graph in $\mathcal{G}(n, 0.001)$.

Diestel[353] has pointed out this often-encountered insensitivity to changes of p. It is a property of 'almost all (random) graphs[5]': for most characteristics, the critical order of magnitude in p that determines the appearance or otherwise of the behaviour is typically a function of the number of nodes, n, that tends to zero as $n \to \infty$. If p is allowed to change with n then for p with order of magnitude under $1/n^2$ a random graph G is, with near certainty, totally disconnected – it has no links at all. As p is raised, the graph begins to gain structure so that from around $p = n^{-3/2}$ components likely exist, having at least two vertices, which become trees as p increases further. The first cycles are seen to appear for $p \sim n^{-1}$, shortly after which the first crossing chords begin to make the sub-graph(s) non-planar so that they can no longer be embedded in a 2-dimensional space (this, of course, provides the theoretical basis for the behaviour exhibited in figure 6.6).

Bollobás[355] discusses the emergence of the 'Giant Component'. Using the definitions that a component is *small* if it has fewer than $\frac{1}{2}n^{2/3}$ vertices and *large* if it has at least $n^{2/3}$ vertices, Bollobás presents and proves the theorem ([355] p142, Theorem 6.8):

> *A.e. graph process* $G = (G_t)_0^N$ *is such that for every* $t \geq t_1 = \lceil n/2 + 2(\log n)^{1/2} n^{2/3} \rceil$ *the graph* G_t *has a unique component of order at least* $n^{2/3}$ *and the other components have at most* $n^{2/3}/2$ *vertices each.*[6]

Aldous and Pittel [356] obtained a consistent result when they considered emergence of the giant component in a non-uniform randomly evolving graph process,

[5] In random graph theory it is recognized that although a system for generating random graphs may appear as though it could deliver many potentially different graphs, this is not the case. Thus it is common to talk about *the* random graph. That almost every two infinite graph random are isomorphic is a theorem that was first proved in Paul Erdös and Alféd Rényi [354].

[6] '*A.e.*' is a graph-theoretic abbreviation meaning 'almost every'.

with vertices immigrating at rate n and each possible edge appearing at rate $1/n$ and found that the detailed picture of emergence of giant components with $O(n^{2/3})$ vertices is the same as in the Erdös - Rényi graph process with the number of vertices fixed at n at the start.

As Diestel puts it, this one component outgrows the others "until it devours them around $p = (\log n)n^{-1}$, making the graph connected." ([353] p.241.) In graph theoretic parlance, the analogy is drawn with evolutionary processes in biology: the properties shared by almost all graphs are considered to be typical, in some sense, of the 'species'. As in nature, random graphs evolve in a manner suggestive of punctuated equilibrium whereby link probabilities exhibit criticality, marking thresholds that separate almost every graph in possession of some property from almost every graph that is so lacking. Real functions $t = t(n) \neq 0$, called *threshold functions* are defined for a graph property P. All non-trivial graph properties that *increase* – for example, *connectedness* – have been shown to have corresponding threshold functions and they are only ever unique up to a multiplicative constant.

The HQS modelling and arguments thus far presented have displayed various facets of, and are entirely compatible with, these features of random graph theory. This will become a recurring theme.

6.10 Concluding remarks for HQS

The concept of a HQS was arrived at by deconstructing the functional integral formulation of quantum field theories, retaining only those aspects that were considered so fundamental as to be intrinsic to the formalism, in the sense of being deeper than

any possible emergent features. Recalling equation 6.7.1 (page 184),

$$< \mathcal{N} \mid \mathcal{N} >= \int \mathcal{D}G e^{-S[G]},$$

being the induced effective action corresponding to a quantum 'field' theory of interacting 3-spaces, the ensuing interpretation in terms of random graph theory has demonstrated the existence of emergent geometrical structures, the gebits, providing intrinsic S^3 characteristics as the most probable form and hence, consistent with the principle of threshold functions, the dominant mode of behaviour.

The next problem for the HQS approach is to demonstrate a natural classical sequencing – whether a HQS can induce the basic phenomenon of time, or at least some notion of what this might mean at such a fundamental level.

Introduce a complete set of functions $\{f_\alpha(G)\}$ for which

$$\delta(G^2 - G^1) = \sum_\alpha f^*_\alpha(G^2) f_\alpha(G^1) \tag{6.10.1}$$

where the superscript on G^i labels copies of $\{G_a\}$. Then, with $S^{(n)}[G] \equiv \frac{1}{n} S[G]$,

$$\begin{aligned} < \mathcal{N} \mid \mathcal{N} > &= \int \mathcal{D}G^2 \int \mathcal{D}G^1 e^{-\frac{1}{2}S[G^2]} \delta(G^1 - G^2) e^{-\frac{1}{2}S[G^1]} \\ &= \sum_\alpha \int \mathcal{D}G^2 1 \cdot e^{-S^{(2)}[G^2]} f^*_\alpha(G^2) \int \mathcal{D}G^1 f_\alpha(G^1) e^{-S^{(2)}[G^1]} \cdot 1, \qquad (6.10.2) \\ C^{(2)}_\alpha &= \int \mathcal{D}G^1 f_\alpha(G^1) e^{-S^{(2)}[G^1]} \cdot 1 \end{aligned}$$

The $C^{(2)}_\alpha$ can be thought of as a transition amplitude, where the '1' represents the beginning of the 'universe', i.e. with all G equally likely. Continue inserting complete

sets

$$C_\alpha^{(2)} = \sum_\beta \int \mathcal{D}G^1 f_\alpha(G^1) e^{-S^{(4)}[G^1]} f_\beta^*(G^1) \int DG^2 f_\beta(G^2) e^{-S^{(4)}[G^2]} \cdot 1$$

$$C_\alpha^{(2)} = \sum_\beta A_{\alpha\beta}^{(4)} C_\beta^{(4)} \tag{6.10.3}$$

$$A_{\alpha\beta}^{(n)} = \int \mathcal{D}G f_\alpha(G) e^{-S^{(n)}[G]} f_\beta^*(G)$$

More insertions give

$$C_\alpha^{(2)} = \sum_{\beta\gamma\ldots} A_{\alpha\beta}^{(n)} A_{\beta\gamma}^{(n)} \ldots C_\cdot^{(n)} \tag{6.10.4}$$

which has the form of an all-inclusive quantum-like multiple sequencing, with the appearance of 'initial' and 'final' states. Figure 6.14 illustrates the idea and the construction is analogous to the form of quantum mechanical sums over intermediate states.

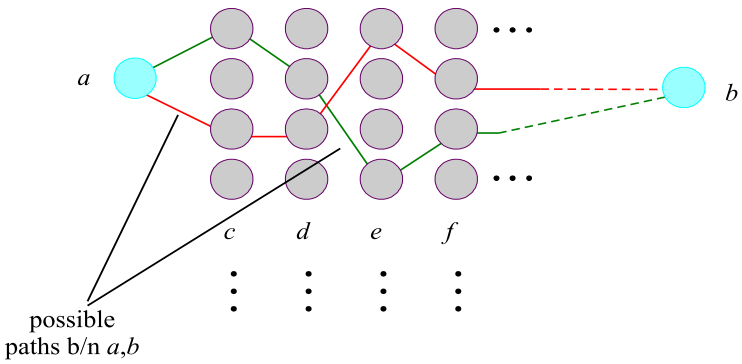


Figure 6.14: Quantum-like multiple sequencing
illustrating 'initial' and 'final' states with some (arbitrary) number of 'intermediate states' between them.

But the time phenomenon is about restricted or classical sequencing, with only some residual quantum phenomena. A unique sequencing could correspond to a *history* of space, defining a space-time block.

A possible macroscopic unique sequencing or history is a partition of $< \mathcal{N} \mid \mathcal{N} >$ such that different histories, by definition, have negligible interference. They are decoherent. They are classical. Hence it must be necessary to look for a particular choice of complete set $\{f_\alpha(G)\}$ for which some of the members generate decoherent and robust histories of any emergent self-synthesizing spatial networks that appear. This has some resemblance to the consistent histories approach to standard quantum theory by Griffiths [357], Omnès [358] and Gell-Mann and Hartle [359].

There is a limit to the usefulness of these complete set insertions – the action $S^{(n)}$ increasingly flattens, so that fluctuations or deviations from the condensate become more extreme, suggesting that any time-like sequencing description has limited relevance at very short time intervals.

At that level of the HQS, it is unreasonable to expect anything other than the usual historical modelling of time along with its deficiencies. The difficulty arises because the deconstruction began with ensembled quantum field theory but it is not possible to recover individuality and actuality from ensembles – a simple fact that has been problematic for quantum theory since its inception. To go beyond this limitation requires that the deconstruction be carried one stage further, incorporating a scheme that both motivates and accommodates changes in the values of the network variables that define the gebit connectivity via the graph links.

Chapter 7

Heraclitean process systems

Change is not merely necessary to life - IT IS LIFE.

– Alvin Toffler

7.1 Introduction

In this chapter, the ideas put forward in describing a HQS are extended by further abstraction and by the introduction of an explicit mechanism to infuse *vitae essentia* into the model. Metaphorically, a HQS provides the 'building blocks' from which a nascent universe is realized. The term 'Heraclitean process system' (HPS) was introduced in [360] to distinguish these further developments from the precursor HQS theory.

Each gebit in a HQS is represented by a tree graph, which is itself an N-vertex subgraph of the greater system – a sparse random graph over M monads defined by the Γ of equation 6.5.8 on page 163, being the multi-element deviations from the condensate. In terms of the whole graph, since $M \to \infty$ the set of monads has

cardinality tending to $\aleph_0$. Because gebits are subsets of M and because they are the most likely form of connected sub-components, it follows that one might expect to find a countable set of gebit trees also with cardinality tending to $\aleph_0$. As large as that set might be, however, it may be regarded for all practical purposes and without loss of generality as being *finite* since, by induction, the mechanism to be described always may be readily extended to accommodate one more gebit.

In the sense that a HQS was concerned with *trees*, a HPS is concerned, in the first instance, with a *forest* comprising all of the monads in M that are components of gebits. The network of connections amongst these monads is described by an adjacency matrix B and since any labels attached to the monads are arbitrary, the disconnected gebit components $g_1, g_2, \ldots, g_n$ may be arranged by a similarity transformation into block diagonal form within B:

$$B = \begin{bmatrix} g_1 = \begin{pmatrix} 0 & 1 & \cdots \\ -1 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} & 0 & 0 & \cdots & 0 \\ 0 & g_2 & 0 & \cdots & 0 \\ 0 & 0 & g_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & g_n \end{bmatrix} \tag{7.1.1}$$

With this forest as a starting point and exploiting the fact that functional integrals can be thought of as arising as ensemble averages of Wiener processes (normally associated with Brownian-type motions in which random processes are used in modelling many-body dynamical systems), HPS theory proceeds from the realization that

the precursor HQS theory failed, ultimately, because it could not produce the phenomenon of time. Although there were suggestions that some notion of temporal behaviour could be recovered, by inference, from evidence of a unique sequencing corresponding to a history of space, this falls far short of recovering time itself. One does not catch a rabbit by spotting its tracks. The reason for the failure of the HQS model lies in its assumption that time is an emergent phenomenon – in §6.1 on page 145, spacetime was explicitly discarded. This was a mistake because there is nothing that is *like* time with which to model time, and the downfall of HQS theory in this regard turns out to be the logical consequence of proceeding from a false assumption, in a sense providing a 'proof' by contradiction that time must be considered truly fundamental.

7.2 HPS theory

7.2.1 Recursion, stochasticity, and process

HPS theory progressed from HQS theory by recognizing the iterative character necessary to model the universe as a self-organizing and self-referential information system (see §5.2 on page 135). Hence, an iterative process is introduced by way of the recursion relation or mapping

$$B_{ij} \to B_{ij} - \alpha(B + B^{-1})_{ij} + w_{ij}, \quad i,j = 1,2,...,2N; \quad N \to \infty, \tag{7.2.1}$$

which henceforth often will be referred to as the '*iterator*', for obvious reasons. For reasons that are not yet obvious, the second term in the iterator will be referred to as the '*binder*'.

As with HQS, the B_{ij} are link variables for relational information shared by two monads i and j, the B_{ij} are real-valued and anti-commuting, so that $B_{ij} = -B_{ji}$ with $B_{ii} = 0$ to preclude self-information. The anticommuting $w_{ij} = -w_{ji}$, also such that $w_{ii} = 0$, are independent random variables for each ij pair and provide the noise of self-referencing, hence, the term '*self-referential noise*' (SRN) described in Part I (in particular, §4.5.2 on page 132). The non-critical parameter α moderates the extent to which the mapping is influenced by the non-linear second term.

7.2.2 HPS as a Wiener process

Equation 7.2.1 has the form of a Wiener process[1] and, as with HQS in §6.1, it arose from considerations founded in the GCM of QCD (see [345] – [350] for reviews). In the GCM, the bilocal-field correlators for meson and baryon correlations are given by the generating functional

$$Z[J] = \int \mathcal{D}B^{\theta} \exp\left(-S[B] + \int dx^4\, d^4y\, B^{\theta}(x,y)\, J^{\theta}(x,y)\right), \tag{7.2.2}$$

where $x, y \in E^4$, the 4-dimensional Euclidean spacetime being chosen because the use of the Euclidean metric is known to select the vacuum state of the quantum field theory and hadronic correlators are required for vacuum-vacuum transitions. Analytic continuation of $x^4 \rightarrow ix_0$ then allows for the Minkowski correlators to be obtained. Integrating out the gluon variables and changing integration variables to go from Grassmannian functional integrations over quarks to bilocal-field functional

[1] A Wiener process, so named in honour of Norbert Wiener, is a continuous-time stochastic process $W(t)$ for $t \geq 0$ with $W(0) = 0$ characterized by the following two conditions: (i) the increment $W(t) - W(s)$ is Gaussian with mean 0 and variance $(t - s)$ for any $0 \leq s < t$, and (ii) increments for nonoverlapping time intervals (say, $W(t) - W(s)$ and $W(u) - W(v)$) are independent. Brownian motion (i.e., random walk with random step sizes) is the most common example of a Wiener process.

integrations leads, approximately, to equation 7.2.2, where the θ-index labels generators of spin, colour and flavour. The form of (7.2.2) is effective for extracting hadronic phenomena because the vacuum state of QCD is analogous to a BCS-like superconductive state, with the $q\overline{q}$ Cooper pairs being described by such non-zero $\overline{B}^{\theta}(x,y)$ as are determined by the Euler-Lagrange equations of the action,

$$\left.\frac{\delta S[B]}{\delta B^{\theta}(x,y)}\right|_{\overline{B}^{\theta}(x,y)} = 0. \tag{7.2.3}$$

This is a non-linear equation for those non-zero bilocal-fields about which the induced effective action for hadronic fields is to be expanded and the BCS-state/constituent-quark effect is revealed by its non-zero solutions.

7.2.3 A stochastic quantization (Parisi-Wu) approach

An alternative to the approach taken in [345] – [350], where (7.2.2) is evaluated approximately as a functional integral, is to utilize the Parisi-Wu *stochastic quantization* procedure [91], which involves the Langevin iterative map

$$B^{\theta}(x,y) \rightarrow B^{\theta}(x,y) - \frac{\delta S[B]}{\delta B^{\theta}(x,y)} + w^{\theta}(x,y). \tag{7.2.4}$$

Here, the $w^{\theta}(x,y)$ are Gaussian random variables with means, $\mu^{\theta} = 0$ (Gaussian to again follow the examples of SQM – see §2.3.2 on page 60, and quantum fluctuations in stochastic gravity – see §2.4.3 on page 82. This is also the simplest idealization of noise, and the only form compatible with the fluctuation dissipation theorem [217]). This map reaches statistical equilibrium after many iterations and, with analytic continuation back to the Minkowski metric, statistical averaging $< B^{\theta}(x,y)B^{\phi}(u,v)\ldots >$ allows the hadronic correlators to be obtained: particularly,

if one writes $B^{\theta}(x,y) = \phi\left(\frac{x+y}{2}\right)\Gamma\left(x-y, \frac{x+y}{2}\right)$ then $\phi(x)$ is a meson field and $\Gamma(x,X)$ is the meson form factor.

The presence of the stochastic term provides the ergodic behaviour that permits the full structure of $S[B]$ to be explored during the iterations (equivalent to integrating over all values of the $B^{\theta}(x,y)$ in (7.2.2)). The complex quantum phenomena involving bound states of constituent quarks embedded in a BCS conducting state are represented by the correlators, $< B^{\theta}(x,y)B^{\phi}(u,v)\ldots >$, with the E^4 Euclidean spacetime playing a passive and completely classical background rôle.

The correspondence between this schemata and the iterative map introduced in (7.2.1) is apparent, and thus the rationale of making that choice for HPS:

- equation 7.2.4 reveals quantum behaviour;
- equation 7.2.1, the iterator, obtains from (7.2.4) by the elimination of the E^4 background (no such prior structure is permitted in a pregeometric theory), removal of indices so that $B^{\theta}(x,y) \rightarrow B_{ij}$, and retention of a simple form for $S[B]$, namely

$$S[B] = \sum_{i>j} B_{ij}^2 - \mathrm{TrLn}(B). \tag{7.2.5}$$

Does the discretized Langevin equation, (7.2.1), continue to exhibit quantum behaviour? More particularly, does it exhibit quantum behaviour while yet retaining, reinforcing, and even exploiting, the dominant-mode behaviour of the HQS gebits – emergent geometrical structure with intrinsic S^3 characteristics to serve as an arena for such quantum behaviour?

7.2.4 The iterator and ensemble averaging

Starting the iterator with $B_{ij} \sim 0$ represents an initial absence of information or order. If α is taken to zero, the non-linear second term is eliminated and the mapping produces only independent and trivial random walks for each B_{ij}. At any iteration, one might examine the B-matrix by coarse-graining in the manner of HQS and performing a similarity transformation to obtain a 'snapshot' forest graph but, although successive iterations might display structure, of sorts, they could not be expected to exhibit any sense of continuity or order. However, as will be demonstrated, the second term introduces a linking of information (hence the term *binder*), by virtue of its global nature via the matrix inverse, B^{-1}, and this, in conjunction with the noise term, leads to the emergence of self-organization and the mapping models a non-local stochastic information system with intrinsic spatial and time-like behaviour while also exhibiting residual non-local and random processes characteristic of quantum effects. Although any specific history of such a noisy iterative process is unique to each realization of the mapping, averages over an ensemble of possible histories can be determined and these have the form of functional integrals. The concept of an ensemble average for any function f of B at iteration $c = 1, 2, 3, \ldots$ is expressed by

$$< f[B] >_c = \int \mathcal{D}B \, f[B] \, \Phi_c[B], \tag{7.2.6}$$

where $\Phi_c[B]$ is the ensemble distribution. The Fokker-Planck equation obtains by the usual construction for Weiner processes:

$$\Phi_{c+1}[B] = \Phi_c[B] - \sum_{ij} \eta \left(\frac{\partial}{\partial B_{ij}} \left(\left(B + B^{-1} \right)_{ij} \Phi_c[B] \right) - \frac{\partial^2}{\partial B_{ij}^2} \Phi_c[B] \right). \tag{7.2.7}$$

In the quasi-stationary regime, it is found that

$$\Phi[B] \sim \exp(-S[B]), \tag{7.2.8}$$

with the action

$$S[B] = \sum_{i>j} B_{ij}^2 - \mathrm{TrLn}(B), \tag{7.2.9}$$

and then the ensemble average is

$$\frac{1}{Z} \int \mathcal{D}B\, f[B]\, \exp(-S[B]), \tag{7.2.10}$$

with Z such that averages are normalized correctly. The ensemble information is sensitive to changes in B_{ji} and that sensitivity is probed by the $(ij)^{\mathrm{th}}$ element of the matrix inverse (providing the connection between (7.2.1) and (7.2.9))

$$\left(B^{-1}\right)_i j = \frac{\partial}{\partial B_{ji}} \mathrm{TrLn}(B) = \frac{\partial}{\partial B_{ji}} \ln \prod_{\alpha} \lambda_\alpha[B], \tag{7.2.11}$$

where the information is in the eigenvalues $\lambda_\alpha[B]$. A further transformation

$$< f[B] > = \frac{1}{Z} \int \mathcal{D}\overline{m}\, \mathcal{D}m \mathcal{D}B\, f[B]\, \exp\left(- \sum_{i>j} B_{ij}^2 + \sum_{i,j} B_{ij} \left(\overline{m}_i m_j - \overline{m}_j m_i\right) \right)$$

$$= \frac{1}{Z} f[\frac{\partial}{\partial J}] \int \mathcal{D}\overline{m}\, \mathcal{D}m\, \exp\left(- \sum_{i>j} \overline{m}_i m_j \overline{m}_j m_i + \sum_{i,j} J_{ij} \left(\overline{m}_i m_j - \overline{m}_j m_i\right) \right) \tag{7.2.12}$$

expresses the ensemble average in terms of Grassmannian fermionic elements in a manner akin to that of HQS theory in §6.1, illustrating that HPS theory both extends and generalizes its predecessor.

7.3 Computational investigations of the iterator

Computational and numerical investigations of the iterator were undertaken with the aim of identifying and exploring explicitly the behaviour of the mapping and the dynamical features it brings to the HPS model. To investigate the general behaviour and characteristics, the following methodology was employed:

1. define the B-matrix at start-up by initializing an $n \times n$ matrix such that all entries are zero except for a '*seed*' comprising a small antisymmetrized gebit submatrix, which is a randomly generated tree graph adjacency matrix.

2. define the SRN matrix at start-up by initializing a $w_{n \times n}$ matrix such that its upper-triangular entries are drawn randomly from a specified distribution and its lower-triangular entries have the corresponding anti-symmetric values.

3. specify a value for the parameter α and a cut-off for the coarse-graining threshold to discriminate 'large' links. Initially, these are rather arbitrary choices to elicit behaviours that are indicative of generic properties and to provide a basis for making subsequent and better informed choices for these parameters.

4. begin iteration. At each step, regenerate the SRN matrix with new random values, update the B-matrix, construct a binary-valued matrix B' by passing B through a step function via the coarse-graining threshold, generate a plot to display the state of the network, and record successive values of the:

   - determinant of B
   - 'action' defined by $S[B] = \sum_{i>j} B_{ij}^2 + \det(B)$,

   absolute values of the complex-valued:

   - principal (i.e. maximum) eigenvalue of B
   - minimum eigenvalue of B

   and also record:

   - the number of components (sub-graphs) in B'

- the size of the largest component of B' and its component vertices

5. after some number of iterations, switch off the noise term and continue, for some further number of iterations, to observe the state of the network in like manner.

6. terminate the iterative procedure and generate plots of the collected data.

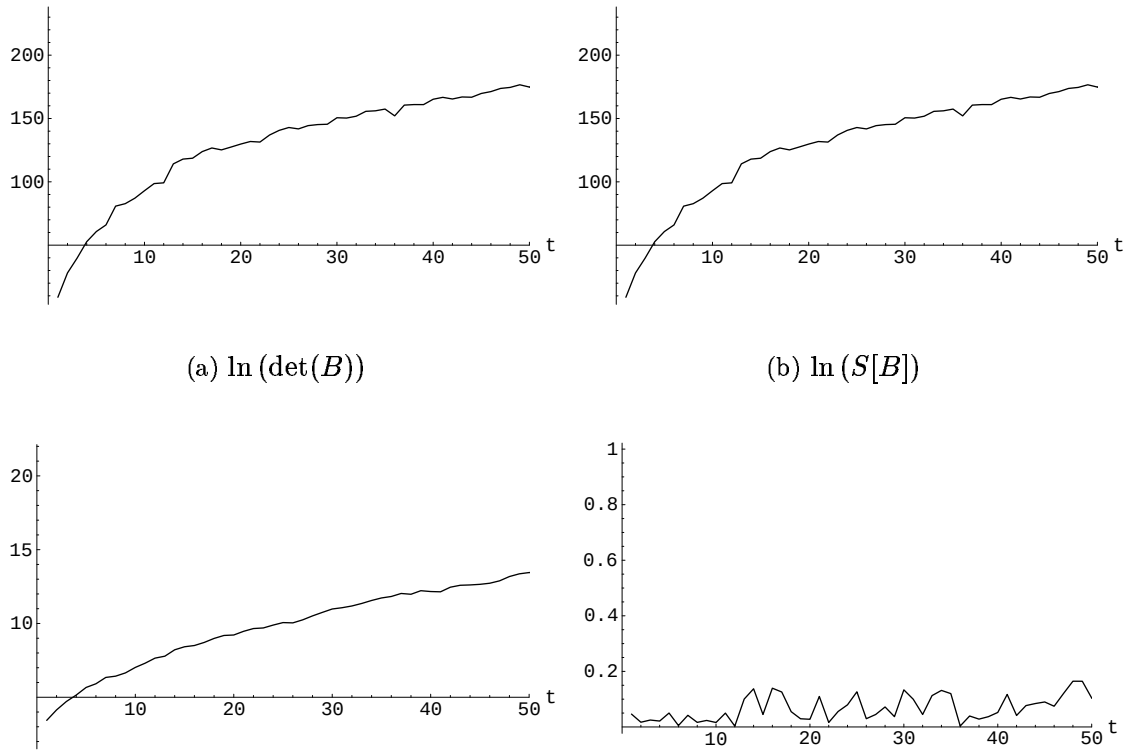
7.3.1 Properties of the iterator

General behaviour and characteristics – Case 1

Because of the intense computational overhead required to conduct these investigations over a reasonable number of iterations, the matrix size was kept small – 120 nodes – following the hypothesis that significant features of the dynamics should not be substantially dependent on scale and thus should be observable even in a model of very modest size. The 'seed' component was a tree over the first six vertices. The SRN component was drawn from a standard normal (i.e. Gaussian) distribution with $\mu = 0$ and $\sigma = 0.1$, the binder parameter is $\alpha = 0.005$, and the coarse-graining cut-off was set to 58% of the maximum B_{ij} at each iteration.

Figures 7.1 and 7.2 report the first fifty iterations comprising the start-up phase for the network during a representative trial of one thousand iterations. Notice first that the least eigenvalue fluctuates close to zero, showing the general stochastic influence of the SRN term. Meanwhile the principal eigenvalue rises steadily as the determinant (represented as $\ln \det(B)$) exhibits power-law growth. Similarly, the action essentially follows the determinant.

Examination of figure 7.2 shows that there is a rapid rise in the number of vertices

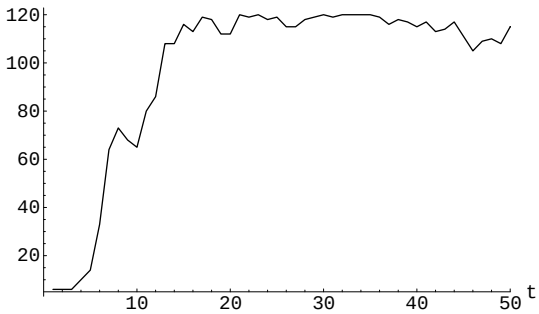


(a) $\ln(\det(B))$ (b) $\ln(S[B])$

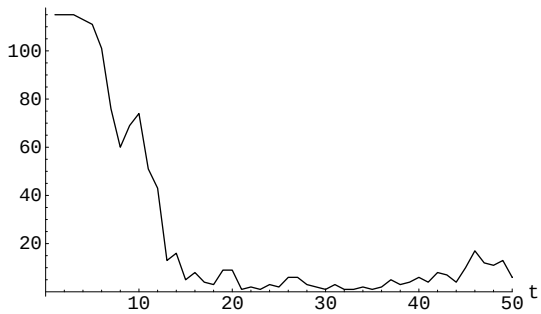
(c) $\max|\text{eigval}(B)|$ (d) $\min|\text{eigval}(B)|$

Figure 7.1: Iterator start-up (i) for $n = 120$
showing the behaviour of the determinant, action, and eigenvalues over the first fifty iterations.

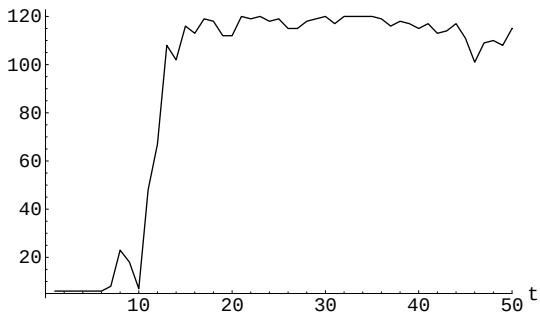
connected by edges whose values lie above the large-link threshold; simultaneously, the number of components drops as connected sub-graphs form and shortly thereafter a large component emerges spontaneously and begins to assert itself, quickly becoming a giant component that dominates the graph by spanning almost all nodes.



(a) Number of nodes with 'strong' links



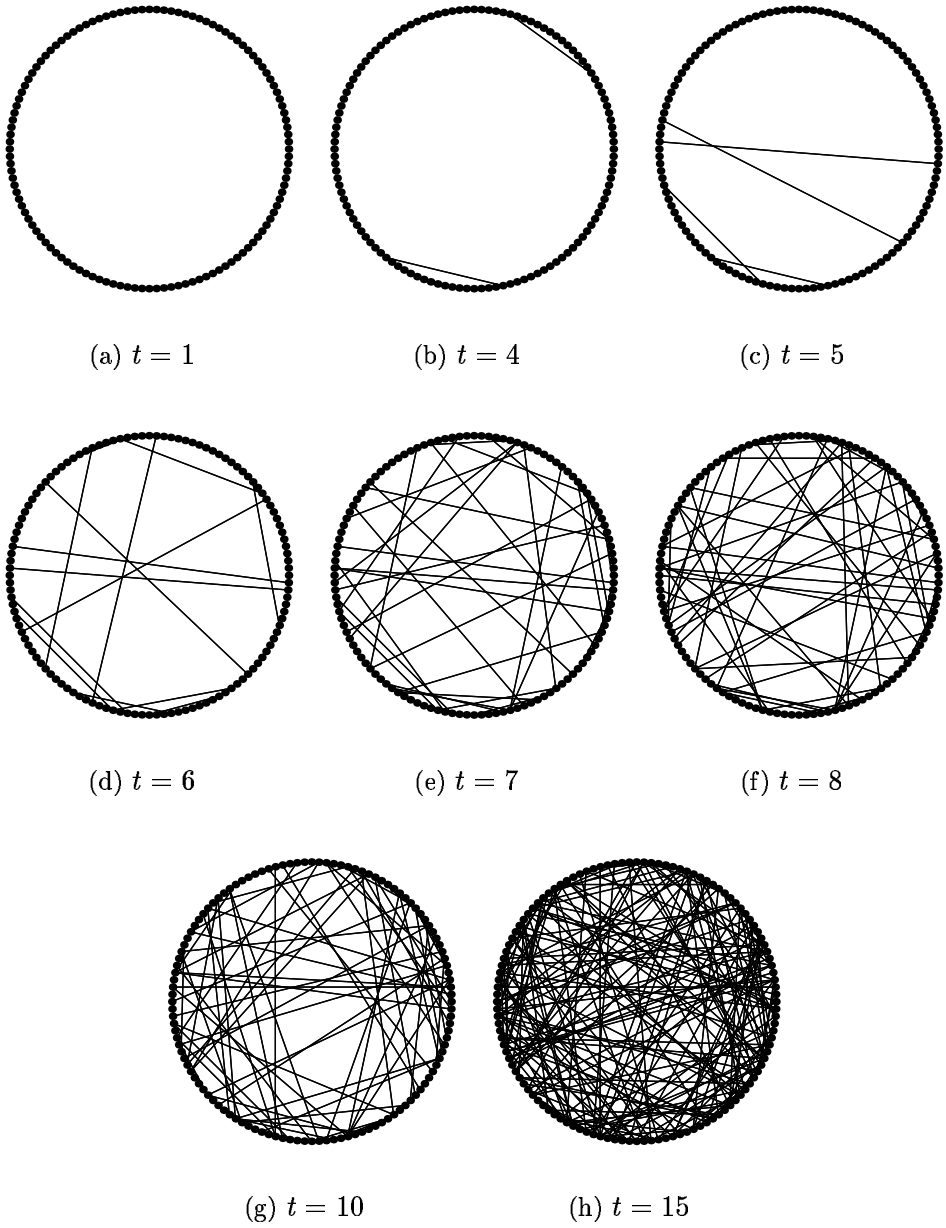
(b) Number of components



(c) Size of largest component

Figure 7.2: Iterator start-up (ii)
demonstrating the rapid appearance of a *giant component* in B' that dominates the graph by spanning most, if not all, vertices.

Figure 7.3 shows circle-embedding representations of B' during the initial part of the start-up phase. At $t = 1$, the graph appears empty because the seed component nodes are adjacent in this representation and no SRN component exceeds threshold. At $t = 4$ the first new edges appear above threshold, one of which survives the next



(a) $t = 1$ (b) $t = 4$ (c) $t = 5$

(d) $t = 6$ (e) $t = 7$ (f) $t = 8$

(g) $t = 10$ (h) $t = 15$

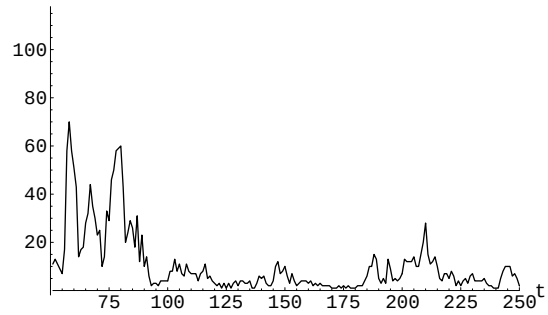
Figure 7.3: Circular embedding representations of B' after coarse-graining showing the development of connected sub-graphs and rapid emergence of the giant component during the 'start-up' phase. Vertices are arranged sequentially on the circumference of a circle; 'large' links – edges with values above the coarse-graining threshold – are shown as chords.

iteration – a connection that persists as yet more edges exceed the threshold, giving rise to connected sub-graphs. Not all edges in these new components show such persistence but many do, and the sub-graphs in turn link up to provide the rapid emergence of the giant component showing in figure 7.3(h) at $t = 15$.

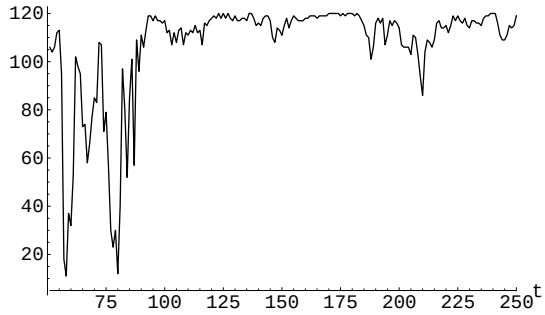
It should be noted, however, that the very rapid growth displayed in this example is somewhat exaggerated by the constraints of the restricted modelling procedure: to elicit the appearance of these behaviours within a relatively small number of iterations, the model is 'over-heated' to some extent by the choices made for α and σ, the former being less and the latter more than one might prefer, though not in the extreme. When taken with the small number of nodes, this leads to the model (like many toy models) being quite volatile and more reactive to fluctuations from stochastic effects than one would expect to occur on any realistic scale. The consequences of these limitations are revealed by the next sequence of iterations.

Figure 7.4 follows the time evolution of the network from $50 < t \leq 250$, showing the number of components (figure 7.4(a)), the size of the largest component (7.4(b)), the number of vertices with edges above threshold (7.4(c)), and the minimum eigenvalue (7.4(d)). Notice that the number of vertices with strong links is high throughout this period – an indication of 'overheating'. The giant component undergoes catastrophic breakdown several times during the first 30 (or so) of these subsequent iterations, becoming re-established by $t = 90$ and remaining quite stable for many iterations thereafter.

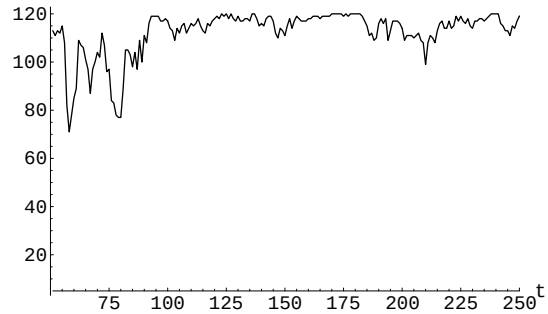
During these events, the minimum eigenvalue fluctuations are consistent with the values seen during start-up, the determinant (figure 7.4(e)) and maximum eigenvalue (7.4(f)) follow the same rising trends as before and, apart from the high number of



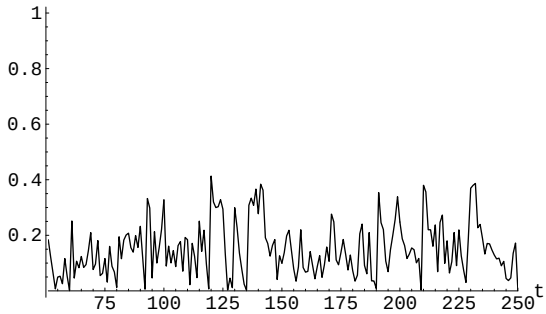
(a) Number of components



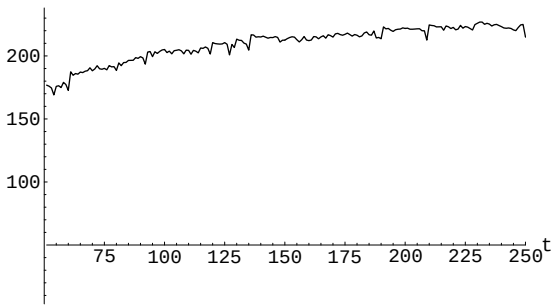
(b) Size of largest component



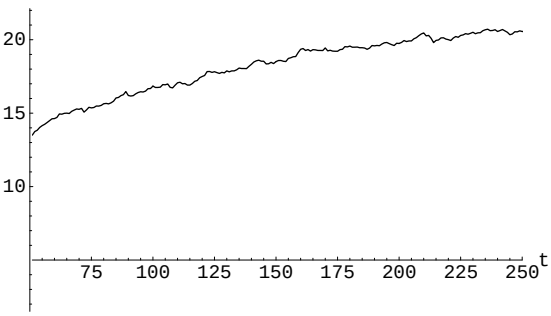
(c) Number of nodes with 'strong' links



(d) $\min |\text{eigval}(B)|$



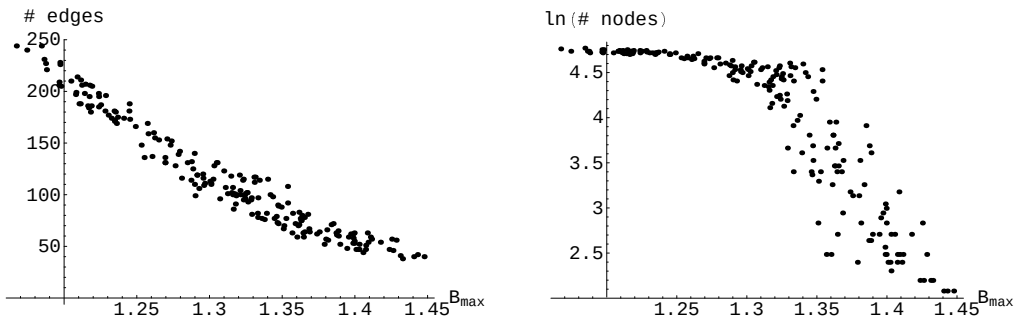
(e) $\ln(\det(B))$



(f) $\max |\text{eigval}(B)|$

Figure 7.4: 'Punctuated equilibrium' or 'avalanches' in the iterator showing catastrophic breakdown of the giant component in B' followed by recovery and relatively long-lived stability.

vertices with links above threshold, there is nothing to indicate the cause of the behaviour. There are, however, two ways in which this can occur and, in each case, the extreme nature of the breakdown is an artefact of the model's limitations. The first reason is that here, with a relatively small graph, just a few critical links dropping below the threshold can be sufficient to disconnect the giant component, causing it to fracture into many parts. One would expect a much larger system to exhibit considerably greater resilience by virtue of its greater capacity to share responsibility for structural integrity; moreover, as the size of the system is increased it becomes increasingly improbable (though not impossible, of course) that the stochastic term will produce simultaneous extreme values sufficient to catastrophically undermine the entire network. The second and most probable reason is the means by which 'large' links are identified. The rare appearance of relatively large values is an important aspect of the model, since this corresponds to having the link probability, p, small – the requirement established in §6.1. Yet, in prescribing a threshold to (artificially) partition the network so as to segregate these large links in order that their relationships may be studied, either a bias is necessarily introduced – yielding the breakdown behaviour – or the full characteristics and consequences of the large link relationships are lost. In this and subsequent investigations, the threshold was chosen simply as a proportion of the maximum value occurring in the matrix at each iteration. This has the result of introducing a bias such that, if a new extremum is substantially greater than its predecessor and also substantially greater than next-largest entries, the new threshold value fails to recapture some, many, most, or even all of the former large-link nodes, bringing about an element of decay and occasional catastrophe to connected components.



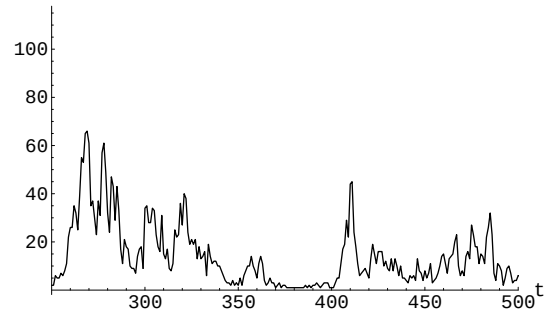
(a) (b)

Figure 7.5: Relationship between $B_{\max}$ and graph components with maximum values of B on the horizontal axis and on the vertical axis: (a) numbers of links above threshold; and (b) log of largest component sizes.

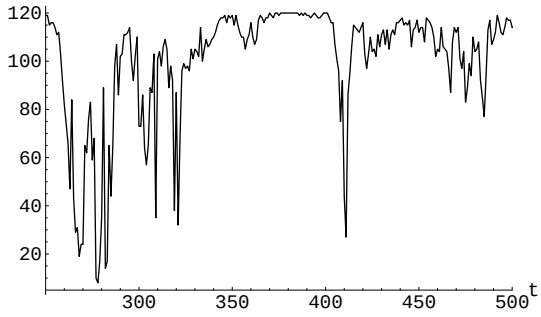
Figure 7.5 shows how the number of links above threshold and the number of vertices in the largest component respond to increasing maxima in B. The bias may be reduced, even eliminated, by adopting some other scheme for identifying large links, such as setting the threshold relative to the median value rather than the maximum; using a moving average of maxima; selecting the k-largest links without regard to their specific values; or taking a sub-maximal cut whenever the maximum is extreme (in a sense, leaving such extreme maxima as members of a higher-order subset much as weaker links might themselves be partitioned further into lower-order subsets). The latter option is appealing because it is suggestive of fractal structure while retaining the notion of the importance of rare large-valued B_{ij}. However, for simplicity (and keeping in mind that this issue resides with the model rather than the theory), it was considered sufficient to proceed as before and leave these remarks as a caveat.

The next 250 iterations demonstrate similar behaviour – relatively long periods of stability in the giant component punctuated by erratic periods peppered with catastrophic events, illustrated by figure 7.6. Here, in particular, the determinant

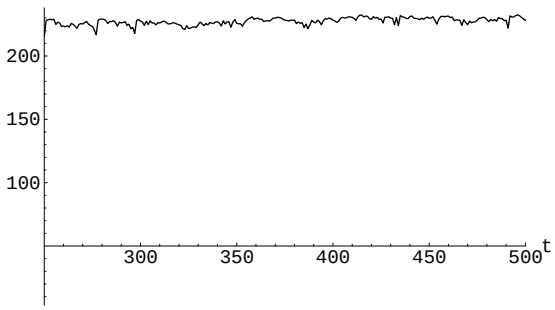
and principal eigenvalue have flattened out, indicating that the network has reached a plateau far from the equilibrium of its initial condensate state.



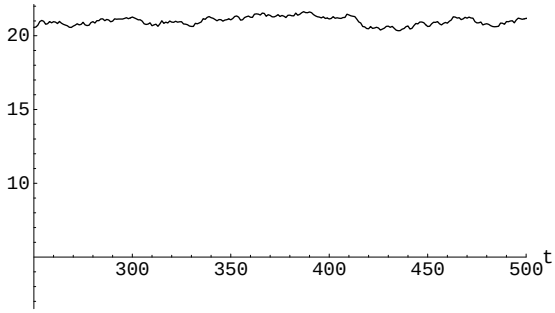
(a) Number of components



(b) Size of largest component



(c) $\ln(\det(B))$

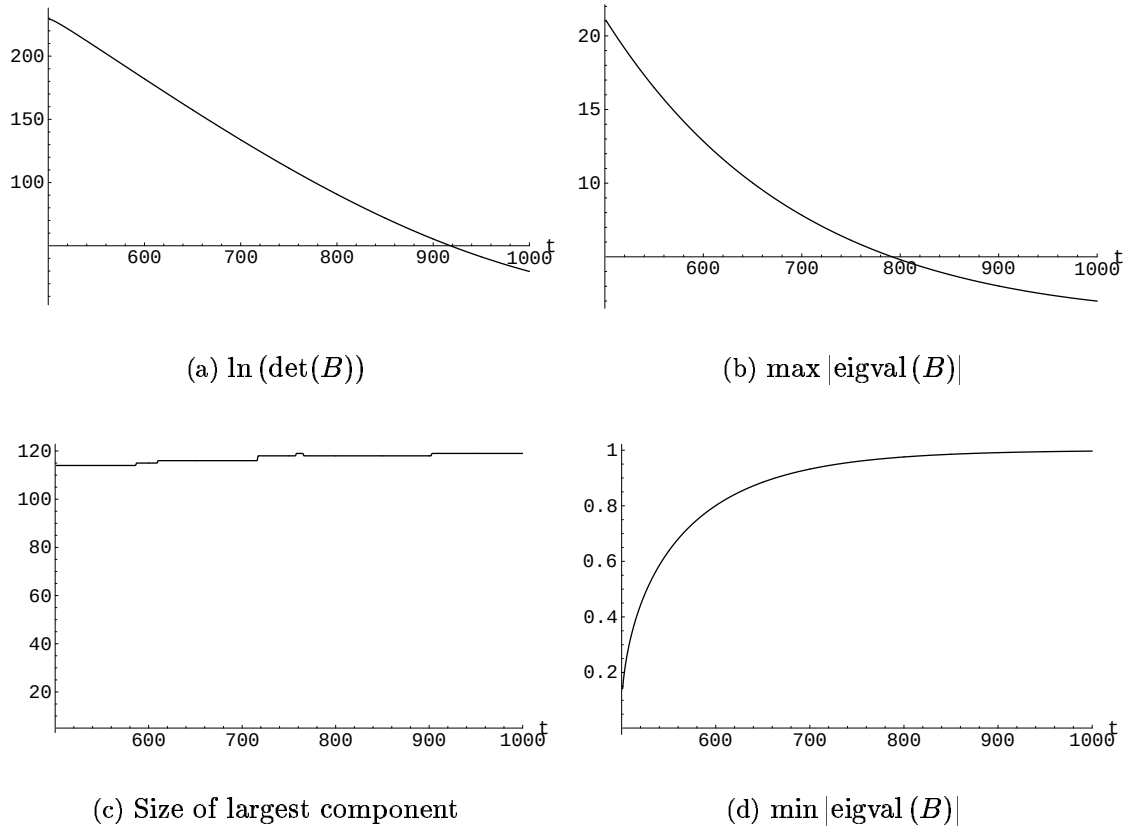


(d) $\max|\text{eigval}(B)|$

Figure 7.6: Longer-term iterator behaviour far from equilibrium showing ongoing 'punctuated equilibrium'. The determinant and principal eigenvalues are relatively flat.

From $t = 500$, the SRN term is switched off and the network allowed to 'cool'. Figure 7.7 on the following page shows the value of the determinant declining smoothly while the principal eigenvalue drops and the least eigenvalue rises as the absolute values of all eigenvalues tend to asymptotic convergence at unity (figure 7.8 on page 224). Meanwhile, the graph remains essentially frozen with the link structure that it possessed when the stochastic input ceased – that is, the relative magnitudes of the edge

variables are maintained while their actual values drop towards zero as the network returns to the condensate.



(a) $\ln(\det(B))$ (b) $\max|\text{eigval}(B)|$

(c) Size of largest component (d) $\min|\text{eigval}(B)|$

Figure 7.7: Iterator behaviour *sans* SRN
showing the system 'cooling' as it returns to the condensate. The giant component remains essentially static.

From its initial emergence during the start-up phase through to $t = 500$ when the SRN term was shut off, the giant component involved every node in the graph. This is shown in figure 7.9. Vertex 28 spent the least number (353) of iterations as a member, while vertex 108 had the highest occupancy rate of 469 per 500 iterations. Across all vertices, the mean occupancy time was 424.4 iterations with a standard deviation of 25.6. During the cooling-off phase from $500 < t \leq 1000$, only six vertices

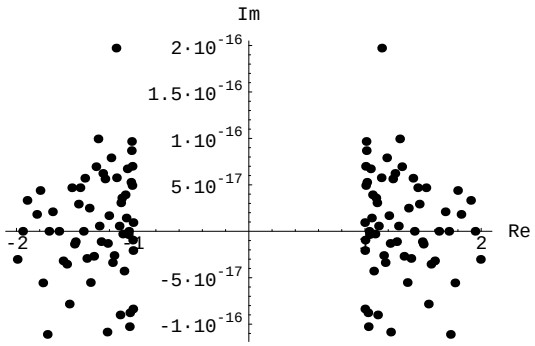
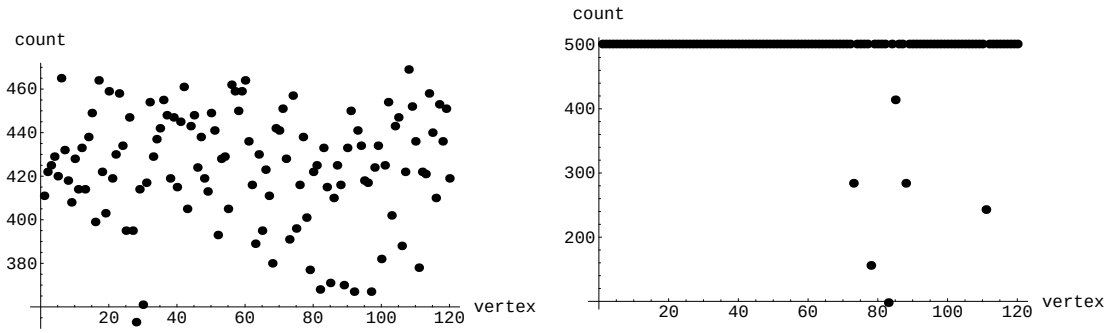


Figure 7.8: Iterator eigenvalues after cooling

spent less than the full period as members of the giant component and these were taken up as the link magnitudes merged into the condensate.



(a) start-up phase $t = 1 \ldots 500$ (b) cooling phase $t = 501 \ldots 1000$

Figure 7.9: Occupancy rates for nodes in the giant component

Using Mathematica's `DiscreteMath'Combinatorica` package, analysis of the B'-graph at $t = 500$ revealed that it consists of one disconnected node together with the giant component containing 119 vertices connected by 222 edges. Within the graph there are 22 cycles ranging in length from 3 to 14 links ($\overline{x} = 7,\ \sigma = 3.2$). The graph depth structure was investigated from the point of view of all nodes to obtain the pair-wise shortest paths between them. The resulting path lengths were counted and the

counts averaged over the 120 vertices, thus giving an averaged set $\{D_k\}$ corresponding to those employed in the previous HQS scheme. Those results were then analyzed using the NonLinearFit algorithm in Mathematica's `Statistics'NonlinearFit` package to obtain the best fit to the parameterized model

$$f(x) = A\sin^2(ax + b) \tag{7.3.1}$$

where $a \approx \frac{\pi}{L}$, with L determined by twice the difference between the maximum path length $k_{\max}$ and k corresponding to $D_{\max}$, and b a phase factor. The resulting best fit is shown in figure 7.10 where the returned parameter values of $A = 37.42$, $a = 0.47$, and $b = -0.144$ give $L = k_{\max} \sim 7$ and $D_{max} \sim 37$.

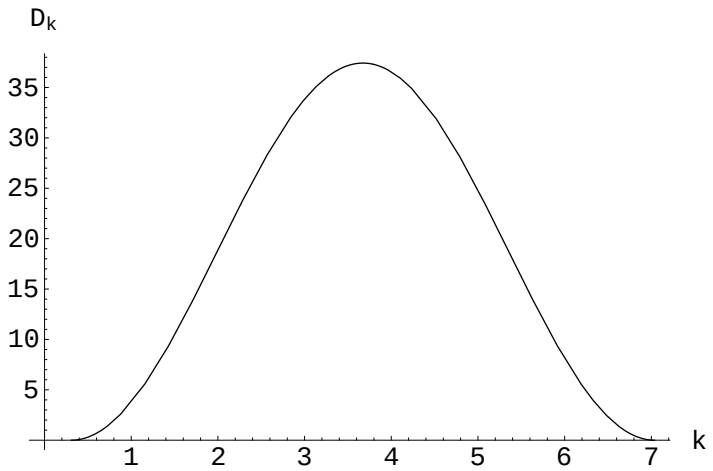


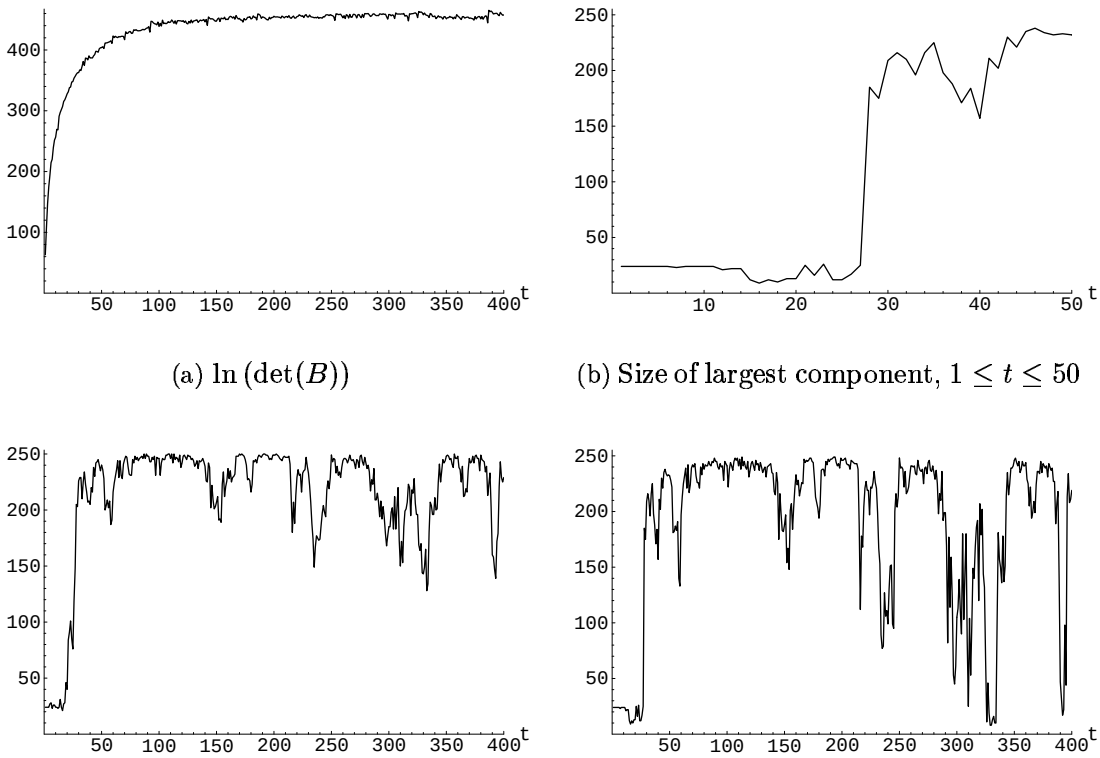
Figure 7.10: Case 1: Nonlinear best fit of B' at $t = 500$.

The $D_{\max}$ and L values do not agree well with the theoretical $D_{\max} = 17$ and $L = 14$ required for consistency with S^3 embedding but they are, however, consistent with the excessive edge count. Taken together with the evidence of catastrophic disruptions noted previously, this offers support to the earlier argument that the model is indeed 'over-heated'.

General behaviour and characteristics – Case 2

If the conjecture is correct that the extreme breakdown of the giant component is an artefact (at least in part) of the model's limitations, a larger construction might be expected to exhibit less extreme behaviour. The procedure is computationally intensive and computer run-times increase exponentially with the size of the model; nonetheless, increasing n from 120 to 250 provides some evidence to affirm the conjecture. Figure 7.11 on the following page shows the results over 400 iterations. Initially, after seeding the graph with a 24-node tree, preliminary 'runs' established that setting the binder amplitude parameter to $\alpha = 0.01$ and $\sigma = \sqrt{0.008}$ allowed the model to achieve equilibrium quickly (while the values of these parameters are non-critical, the competing rôles of their respective terms in the iterator mean that their relative magnitudes establish the 'thermal limit' for equilibrium away from the condensate), substantially reducing the computation time at start-up. The coarse-graining threshold was set at 58% of $B_{\max}$, as before.

Figure 7.11(a) shows the determinant flattening out within the first 100 iterations, while 7.11(b) shows the giant component appearing within the first 30 iterations. In figure 7.11(c), the number of vertices with edges above threshold is consistently high following the formation of the giant component but more variable than previously. Figure 7.11(d) indicates that although there are intermittent failures of the giant component, these tend to be less frequent and (mostly) not as extreme as before. While not conclusive, this result does indicate that the conjecture is not unfounded.



(a) $\ln(\det(B))$

(b) Size of largest component, $1 \leq t \leq 50$

(c) Number of nodes with 'strong' links

(d) Size of largest component

Figure 7.11: Longer-term iterator behaviour far from equilibrium showing ongoing 'punctuated equilibrium'. The determinant and principal eigenvalues are relatively flat.

General behaviour and characteristics – Case 3

Aside from the threshold issues noted earlier (page 220), the matter of determining an appropriate threshold value for the coarse-graining is problematic in itself. If the threshold is too high the resulting graph will contain insufficient edges, relative to the number of nodes, to exhibit significant structure formation. On the other hand, a threshold that is too low will yield a graph that is too dense in its edge count, contrary to the original requirement that the link probability be small. Again, the need to set a threshold resides with the model, *not* with the theory, and is necessitated by the limitations of the modelling procedure. If n is sufficiently large, the threshold can be arbitrarily high (short of totally inhibiting links between monads) since connected components must arise eventually and at some stage thereafter these too will become linked. For the purposes of toy modelling, though, pragmatic considerations take precedence and demand a practical solution. The 58% cut-off used in the previous examples was chosen empirically from a number of trials as a first approximation. In this next example, a model with $n = 200$ was run at the 58% threshold until formation of the giant component, whereupon the cut-off was varied dynamically so as to keep the difference between n and the number of edges within 12% of n. The mapping approaches its equilibrium plateau in figures 7.12(a)–(c) on page 232, while (e) shows the onset of 'tuning' at $t = 22$, with the consequent effects on B' appearing in the numbers of components (d) and number of nodes with 'strong' links (f).

Figure 7.13 on page 233 shows the next 200 iterations, i.e. $50 < t \leq 250$, with the cut-off threshold set at 59.9%. Here, the giant component dominates, yet only occasionally does it include all 200 vertices. As before, the behaviour is punctuated by intermittent collapses, though only one, beginning at $t \sim 150$, might be considered

catastrophic. During this phase, the network continued to settle towards the equilibrium plateau, but now at a slower rate, as indicated by figures (a) - (c). At $t = 250$, investigation of the large-link structure provided the results summarized in figure 7.14 on page 234. The giant component comprises a little over three-quarters of the available vertices with connectivity some 16% higher than that of a tree-graph, leading to the presence of eight cycles – so this sub-graph has the characteristics of the *augmented trees* introduced in §6.1. The nonlinear best fit to $f(x) = A\sin^2(ax+b)$ shows excellent agreement with the predicted S^3 signature, though with some asymmetry in the tails because the the depth structure was obtained using only a simple averaging procedure akin to that first employed for HQS rather than the more sophisticated normalization approach (see page 194).

Continuing the same 'run' for a further 200 iterations, $250 < t \leq 450$ provided the results shown in figure 7.15 on page 235 and the large-link analysis conducted at $t = 450$ is summarized in figure 7.16 on page 236. Figures 7.15(a) - (c) show that the equilibrium plateau has almost been attained since, on average, the respective quantities increase only slightly, while (d) and (e) display a good degree of stability in the giant component. There is a major, yet non-catastrophic, collapse event occurring at $t \sim 360$ but with almost immediate recovery attempts punctuated by lesser collapses suggesting the presence within the network of competing factors. The large-link analysis procedure is computationally too intensive to be conducted at every iteration but one likely, albeit speculative, hypothesis is that these may have been associated with a secondary large component or that critical links were being only weakly restored. The nonlinear best fit analysis (figure 7.16) gives quite good agreement with theory, although the D_k distribution is pinched and $D_{\max} = 32$ much

greater than the theoretical value of 24 due to the greater connectivity (20% more than a tree-graph).

Results for $450 < t \leq 650$ are presented in figure 7.17 on page 237. The network is well settled and now very close to equilibrium (the continuing slight rise in the principal eigenvalue shown in (c) indicates that 'heating' is still occurring but the very small slope of the graph shows that this is very gentle). The behaviour of the giant component and the number of components displays the same characteristics as the previous set of results except that, on average, the size of the largest component is slightly less than before. This is because the cut-off threshold here was raised slightly to 61.9 to compensate for the graph not yet having reached the equilibrium plateau. Figure 7.18 on page 238 displays the results for the nonlinear best fit, which is in excellent agreement with the predicted behaviour (note $D_{\max} = 25$ and $L = 16$ with the respective theoretical values of 24 and 15) and the connectivity is 17% higher than that of a tree-graph with the same number of vertices.

Further iterations over $650 < t \leq 850$ are reported in figure 7.19 on page 239. The behaviour of the principal eigenvalue (figure 7.19(c)) through these iterations demonstrates that the equilibrium plateau has been attained, while figure 7.19(e) shows that the giant component is very robust, with no catastrophic events, yet still punctuated with episodic small collapses. The nonlinear best fit (figure 7.20 on page 240) is in excellent agreement with the HPS theory, with relative connectivity $\sim$15%.

Finally for this example, figure 7.21 on page 241 shows the behaviour over $850 < t \leq 1050$. The determinant and principal eigenvalue have average slope ~ 0 and $B_{\max}$ fluctuates moderately about a constant value. The giant component suffers several collapses, two of which are quite long-lived with onset at $t \sim 865$ and $t \sim 975$,

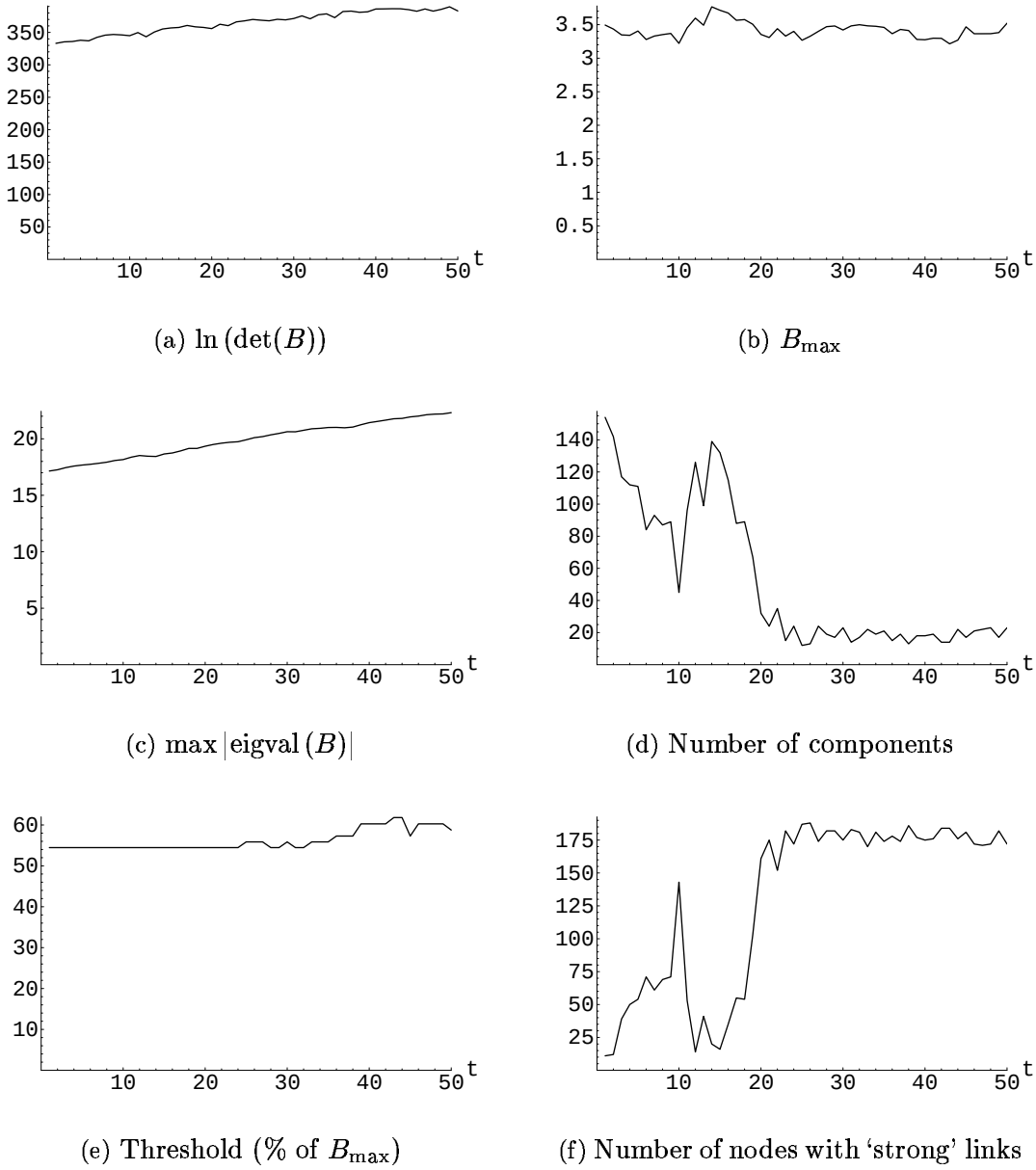
respectively, and, metaphorically, it is tempting to regard the behaviour as analogous to that of convection cells in boiling water. Figure 7.22 on page 242 shows the strong-link analysis, with relative connectivity $\sim$11%, again providing a nonlinear best fit that is in sound agreement with the theory.

General behaviour and characteristics – Case 4

Figures 7.23, 7.24, 7.25, and 7.26 on the following pages show the results for a model with 500 nodes, taken over 800 iterations from start-up. The behaviour of the determinant, $B_{\max}$, principal eigenvalue, number of components, and number of vertices in the largest component are entirely consistent with prior results. Here, the 'seed' tree comprised 50 nodes (figure 7.23(h)) but this immediately decayed prior to the appearance of the giant component at $t \sim 50$, which persisted for the next 500 iterations before suffering its first catastrophic breakdown.

The large-link analyses at $t = 400$ and $t = 800$ were undertaken with cut-off values of 61.65% and 61.76% respectively, consistent with previous values, suggesting that the threshold (and thus the link probability, p) may be independent of n. Again, the nonlinear best fit results are in sound agreement with theory and the relative connectivity of the giant component at the iteration steps shown in figures 7.24 and 7.26 is approximately 13% and 11% respectively.

In this trial, the rôle of the binder was subjected to closer scrutiny by recording at each iteration the determinant, maximum value, and principal eigenvalue of the matrix inverse, B^{-1} (shown in parts (b), (d) and (f) of figures 7.23 and 7.25). The feature of particular interest is the sporadic appearance of very large values in $B^{-1}_{\max}$ and the corresponding principal eigenvalue, behaviour that is highly significant,



(a) $\ln(\det(B))$

(b) $B_{\max}$

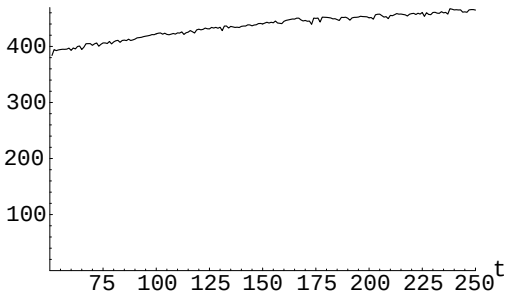
(c) $\max|\text{eigval}(B)|$

(d) Number of components

(e) Threshold (% of $B_{\max}$)

(f) Number of nodes with 'strong' links

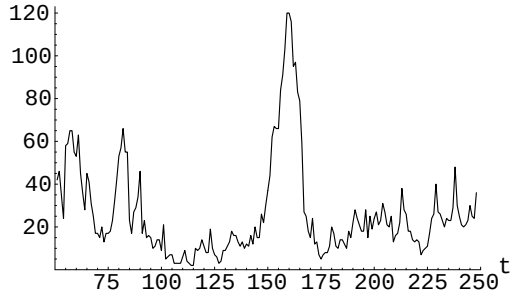
Figure 7.12: 'Tuning' the large-link cut-off threshold
for B-matrix with $n = 200$ over 50 iterations. The onset of 'tuning' is shown in (e).



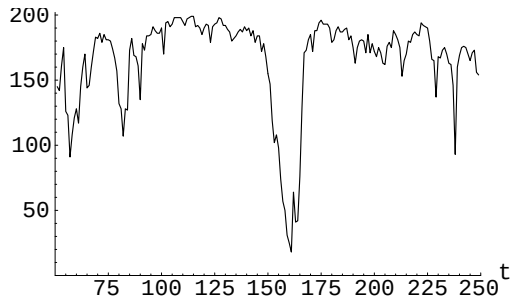
(a) $\ln(\det(B))$

(b) $B_{\max}$

(c) $\max|\text{eigval}(B)|$



(d) Number of components



(e) # vertices in largest component

Figure 7.13: Iterator sample run: $n = 200$, $50 < t \leq 250$

```
{BestFitParameters → {amp → 24.2648, Aa → 0.215604, Bb → -0.0337642}, ParameterCITable →
          Estimate        Asymptotic SE      CI
 amp      24.2648         0.622408           {22.9201, 25.6094}
 Aa       0.215604        0.00527087         {0.204217, 0.226991}   ,
 Bb       -0.0337642      0.0452985          {-0.131626, 0.0640973}
EstimatedVariance → 1.66231,
                                  DF      SumOfSq      MeanSq
             Model                3       3220.35      1073.45
ANOVATable → Error                13      21.61        1.66231,
             Uncorrected Total    16      3241.96
             Corrected Total      15      1239.4
                                    1.        0.464156  -0.422175
AsymptoticCorrelationMatrix → ( 0.464156      1.        -0.904656 ),
                                -0.422175  -0.904656     1.
                                                    Curvature
FitCurvatureTable → Max Intrinsic                   0.205478
                    Max Parameter-Effects           0.0590326 }
                    95. % Confidence Region         0.541488
```

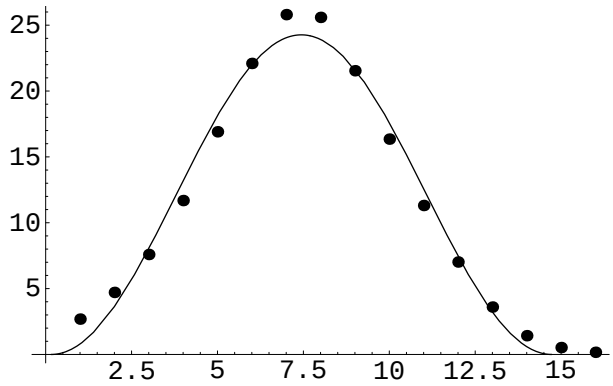
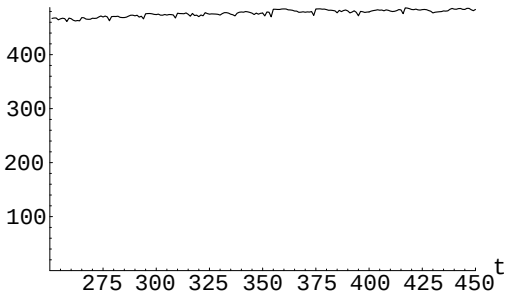
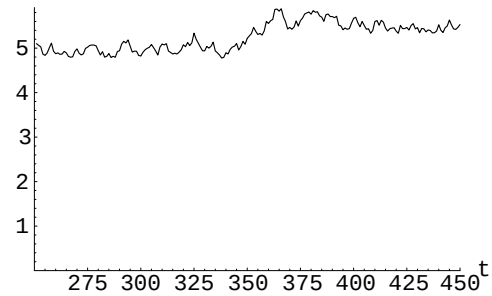


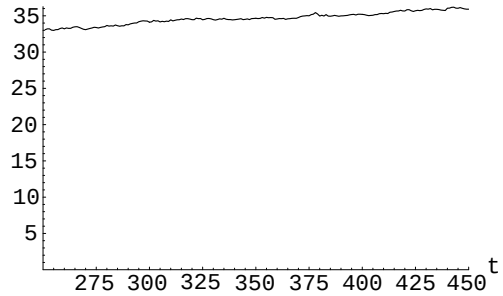
Figure 7.14: Case 3: Nonlinear best fit on B' at $t = 250$ with cut-off fixed at 59.9%. *Mathematica* output for the non-linear best fit algorithm on the giant component, which consists of 155 nodes, 179 edges, and 8 cycles with lengths $\{4, 5, 5, 9, 11, 12, 13, 16\}$. Points are the respective D_k counts fitted to $f(x) = A\sin^2(ax + b)$, from equation 7.3.1, with $A = 24.3$, $a = 0.2$, and $b = -0.03$. Particularly, compare $D_{\max} = 26$ and $L = 16$ with the theoretical values of 24 and 15, respectively.



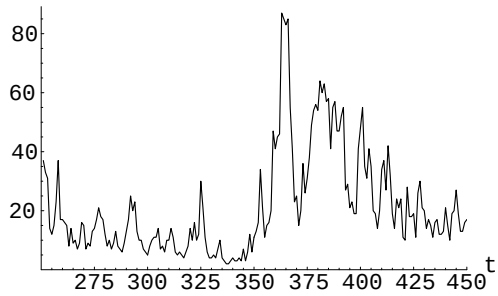
(a) $\ln(\det(B))$



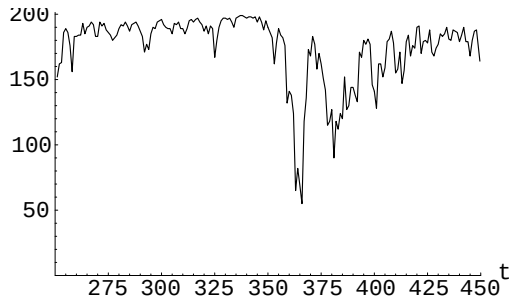
(b) $B_{\max}$



(c) $\max|\text{eigval}(B)|$



(d) Number of components



(e) # vertices in largest component

Figure 7.15: Iterator sample run: $n = 200$, $250 < t \leq 450$

```
{BestFitParameters → {amp → 28.1304, Aa → 0.219602, Bb → 0.119993},

                                Estimate     Asymptotic SE     CI
ParameterCITable →  amp         28.1304      1.4399            {25.0196, 31.2411}
                    Aa          0.219602     0.00909812        {0.199947, 0.239258}   ,
                    Bb          0.119993     0.0812084         {-0.0554472, 0.295433}

EstimatedVariance → 9.92027,

                                    DF      SumOfSq      MeanSq
                    Model           3       4288.56      1429.52
ANOVATable →        Error           13      128.963      9.92027,
                    Uncorrected Total       16      4417.53
                    Corrected Total         15      2040.96

                                                  1.        0.340808   -0.313135
AsymptoticCorrelationMatrix →                  0.340808       1.      -0.874234
                                               -0.313135   -0.874234      1.

                                                                Curvature
FitCurvatureTable →  Max Intrinsic                              0.26078
                     Max Parameter-Effects                      0.184992   }
                     95. % Confidence Region                    0.541488
```

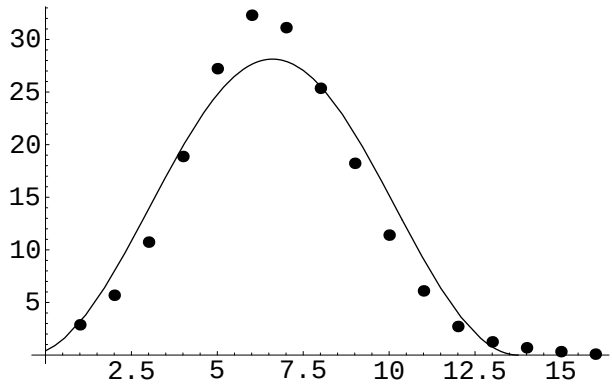
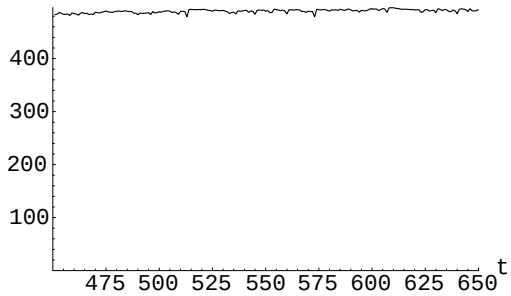
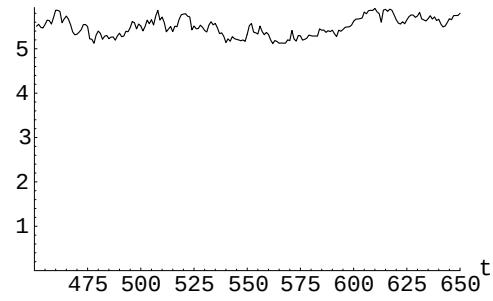


Figure 7.16: Case 3: Nonlinear best fit on B' at $t = 450$

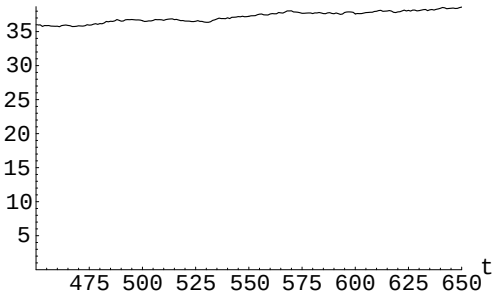
Here, the giant component has 163 nodes, 195 edges, and 9 cycles with lengths $\{3, 3, 4, 5, 5, 8, 12, 16, 18\}$. Points are the respective D_k counts fitted to $f(x) = A\sin^2(ax + b)$, as before, with $A = 28.1$, $a = 0.2$, and $b = 0.1$.



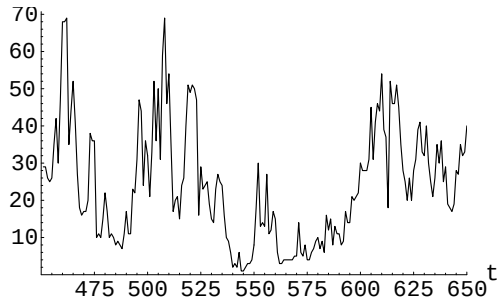
(a) $\ln(\det(B))$



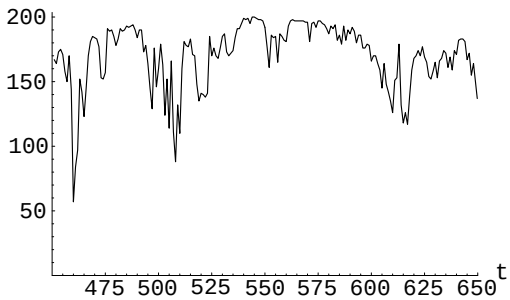
(b) $B_{\max}$



(c) $\max|\text{eigval}(B)|$



(d) Number of components



(e) # vertices in largest component

Figure 7.17: Iterator sample run: $n = 200$, $450 < t \leq 650$

```
{BestFitParameters → {amp → 23.0935, Aa → 0.214886, Bb → 0.0981401}, ParameterCITable →
          Estimate       Asymptotic SE      CI
 amp      23.0935        0.761196           {21.449, 24.7379}
 Aa       0.214886       0.00622133         {0.201446, 0.228326}   ,
 Bb       0.0981401      0.0537138          {-0.0179015, 0.214182}
EstimatedVariance → 2.66879,
                                  DF      SumOfSq      MeanSq
            Model                 3       2935.99      978.663
ANOVATable → Error                13      34.6943      2.66879,
            Uncorrected Total     16      2970.68
            Corrected Total       15      1227.62
                                    1.        0.403851   -0.36416
AsymptoticCorrelationMatrix → ( 0.403851      1.        -0.883333 ),
                                -0.36416   -0.883333      1.
                                                        Curvature
FitCurvatureTable → Max Intrinsic                       0.212589
                    Max Parameter-Effects               0.121462 }
                    95. % Confidence Region             0.541488
```

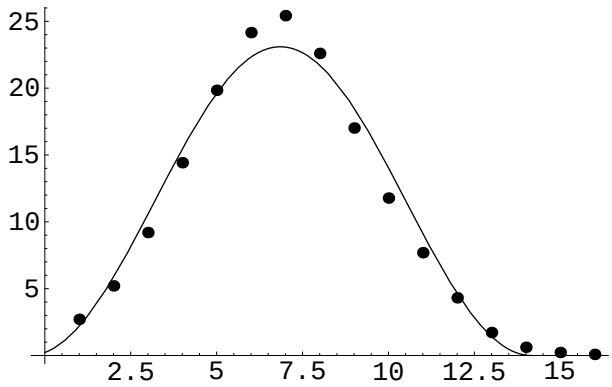
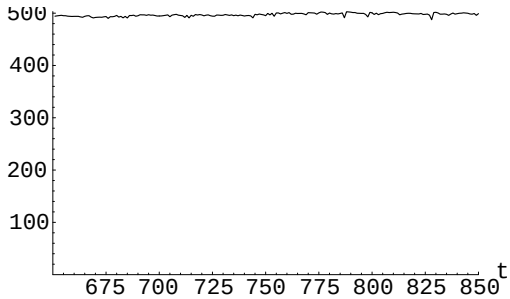
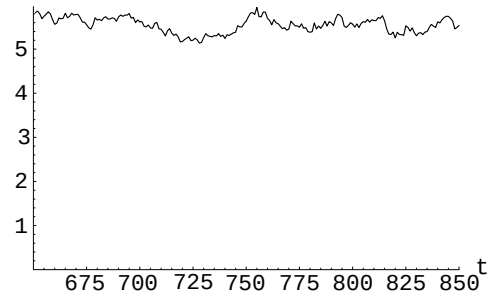


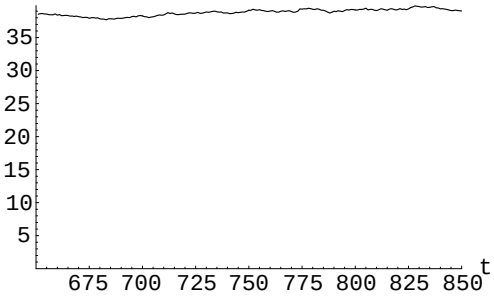
Figure 7.18: Case 3: Nonlinear best fit on B' at $t = 650$ with cut-off fixed at 61.9%. Now the giant component has 144 nodes, 167 edges, and 7 cycles with lengths $\{4, 4, 7, 8, 8, 18, 18\}$. Again, points are the respective D_k counts fitted to $f(x) = A \sin^2(ax + b)$ with $A = 23.0$, $a = 0.2$, and $b = 0.1$. Compare $D_{\max} = 25$ and $L = 16$ with the theoretical values of 24 and 15, respectively.



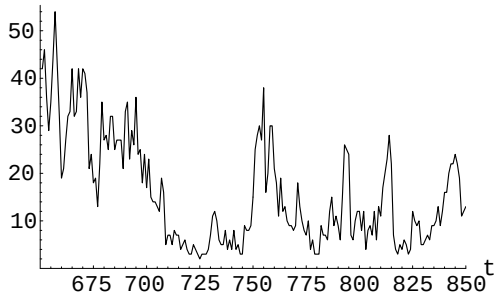
(a) $\ln(\det(B))$



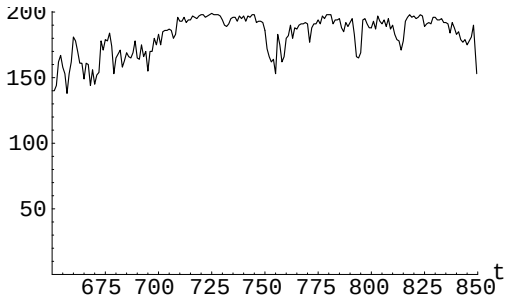
(b) $B_{\max}$



(c) $\max|\text{eigval}(B)|$



(d) Number of components



(e) # vertices in largest component

Figure 7.19: Iterator sample run: $n = 200$, $650 < t \leq 850$

```
{BestFitParameters → {amp → 24.4831, Aa → 0.223796, Bb → -0.00424321},
 ParameterCITable →
          Estimate         Asymptotic SE      CI
  amp     24.4831          0.499717           {23.3943, 25.5719}
  Aa      0.223796         0.00450596         {0.213978, 0.233614}'
  Bb      -0.00424321      0.0359631          {-0.0826, 0.0741136}
 EstimatedVariance → 1.00277,
                                  DF      SumOfSq      MeanSq
              Model               3       3156.61      1052.2
 ANOVATable → Error               12      12.0332      1.00277,
              Uncorrected Total   15      3168.64
              Corrected Total     14      1150.24
                                    1.        0.48729   -0.442343
 AsymptoticCorrelationMatrix → (  0.48729      1.      -0.905536  ),
                                 -0.442343  -0.905536     1.
                                                    Curvature
 FitCurvatureTable → Max Intrinsic                  0.171042
                     Max Parameter-Effects          0.046293  }
                     95. % Confidence Region        0.535265
```

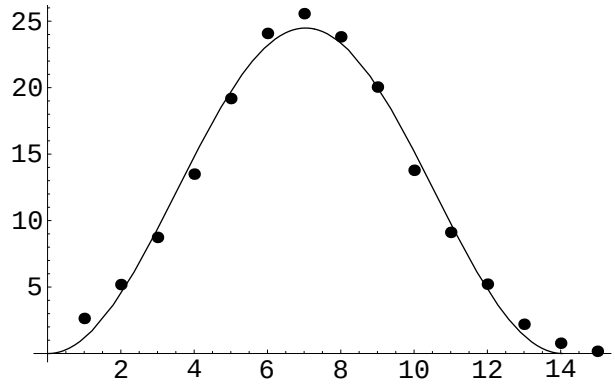
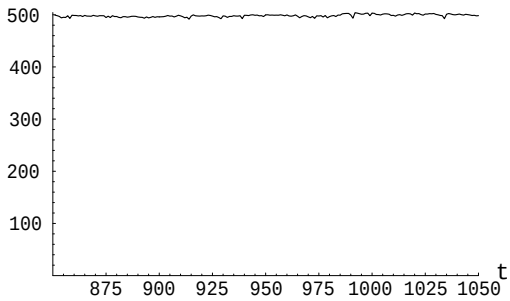


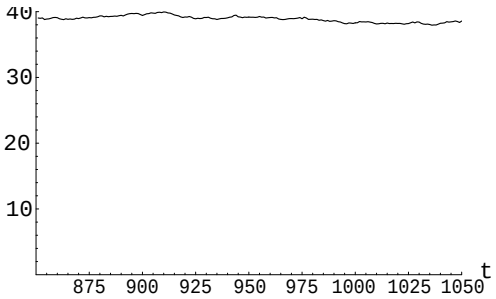
Figure 7.20: Case 3: Nonlinear best fit on B' at $t = 850$

Here, the giant component has 152 nodes, 174 edges, and 5 cycles with lengths $\{3, 9, 9, 12, 13\}$. Points are the respective D_k counts fitted to $f(x) = A\sin^2(ax + b)$ with $A = 24.5$, $a = 0.2$, and $b = 0.0$. Compare $D_{\max} = 25$ and $L = 16$ with the respective theoretical values of 24 and 15.

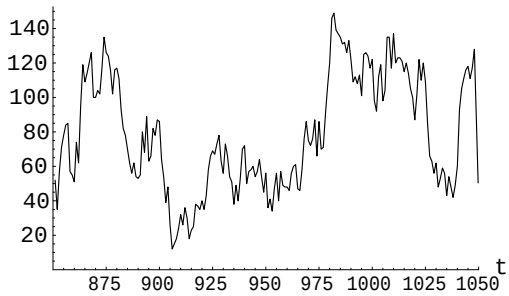


(a) $\ln(\det(B))$

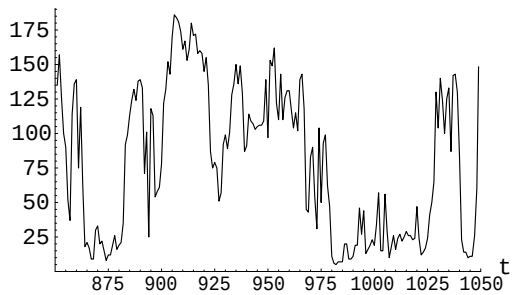
(b) $B_{\max}$



(c) $\max|\text{eigval}(B)|$



(d) Number of components



(e) # vertices in largest component

Figure 7.21: Iterator sample run: $n = 200$, $850 < t \leq 1050$

```
{BestFitParameters → {amp → 22.7787, Aa → 0.216273, Bb → -0.0143769}, ParameterCITable →
          Estimate        Asymptotic SE      CI
 amp      22.7787         0.531524           {21.6304, 23.927}
 Aa       0.216273        0.00471472         {0.206087, 0.226459}   ,
 Bb       -0.0143769      0.0405584          {-0.101998, 0.0732443}
EstimatedVariance → 1.22854,
                                  DF      SumOfSq      MeanSq
             Model                3       2830.62      943.541
ANOVATable → Error                13      15.971       1.22854,
             Uncorrected Total    16      2846.59
             Corrected Total      15      1124.34
                                    1.        0.450374  -0.408824
AsymptoticCorrelationMatrix → ( 0.450374      1.        -0.900662 ),
                                -0.408824  -0.900662     1.
                                                    Curvature
FitCurvatureTable → Max Intrinsic                   0.179174
                    Max Parameter-Effects           0.0755519}
                    95. % Confidence Region         0.541488
```

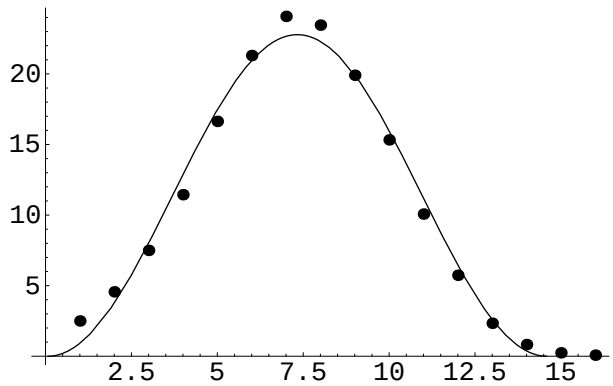


Figure 7.22: Case 3: Nonlinear best fit on B' at $t = 1050$

The giant component now has 149 nodes, 166 edges, and 6 cycles with lengths $\{3, 8, 8, 8, 20, 23\}$. Points are the respective D_k counts fitted to $f(x) = A\sin^2(ax+b)$ with $A = 22.8$, $a = 0.2$, and $b = 0.0$. Compare $D_{\max} = 24$ and $L = 16$ with the respective theoretical values of 24 and 15.

warranting closer investigation in its own right.

7.3.2 The rôle and attributes of the binder

Degeneracy and singular values

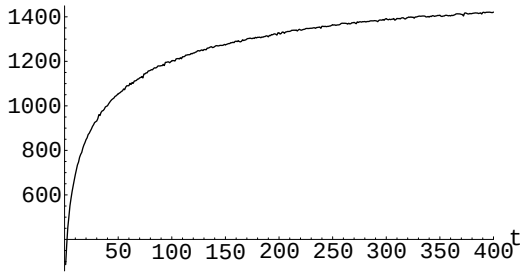
If the intermittent large values in B^{-1} are sufficiently large, their product with the binder strength parameter, α, reinforces otherwise weak links and introduces new links at many locations across the network. To see how this comes about, consider a small tree graph such as the 6-vertex example in figure 7.27 on page 248.

Most tree-graph adjacency matrices are singular (their determinant vanishes) and thus their null space, or *kernel* is a non-empty set and such is the case with this example. The adjacency matrix in figure 7.27(a) is identical in rows (and columns) 2, 3 and 5 and has null space vectors $\{0,-1,0,0,1,0\}$ and $\{0,-1,1,0,0,0\}$, where the non-zero elements identify the nodes associated with the singular values – that is, vertices 2, 3, and 5 in this example. There are very many ways in which redundant entries may arise in the adjacency matrices of graphs, rendering them singular, and for tree graphs (and augmented trees) probably the most common situation is the one depicted here, where a parent node to leaf nodes has degree greater than two[2].

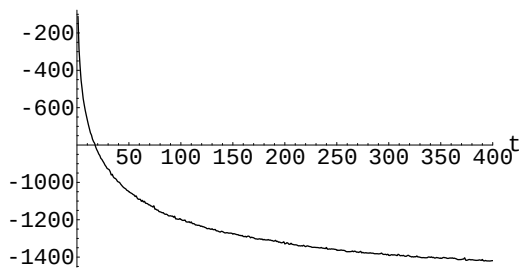
A similar, yet even simpler, example is that of figure 7.28 on page 249. For this graph, the adjacency matrix (in anti-symmetric form as the generator of the tree) and its inverse are given by

$$T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ 0 & -1 & 0 & \delta \\ 0 & -1 & -\delta & 0 \end{pmatrix}; \qquad T^{-1} = \begin{pmatrix} 0 & -1 & \frac{1}{\delta} & \frac{-1}{\delta} \\ 1 & 0 & 1 & 1 \\ \frac{-1}{\delta} & 0 & 0 & \frac{-1}{\delta} \\ \frac{1}{\delta} & 0 & \frac{1}{\delta} & 0 \end{pmatrix} \tag{7.3.2}$$

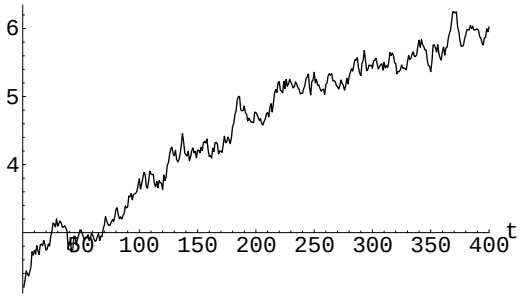
[2]This is an important point to which further attention will be directed, particularly in §8.6.



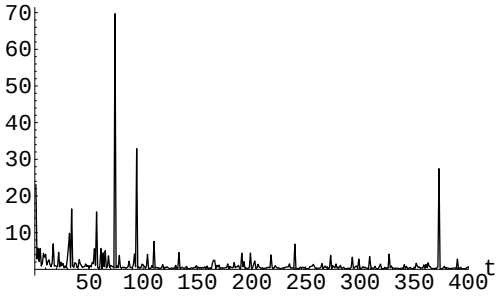
(a) $\ln(\det(B))$



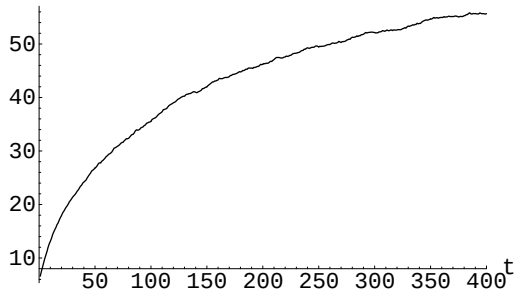
(b) $\ln\left(\det(B^{-1})\right)$



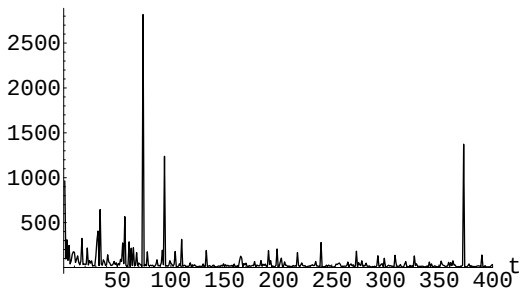
(c) $B_{\max}$



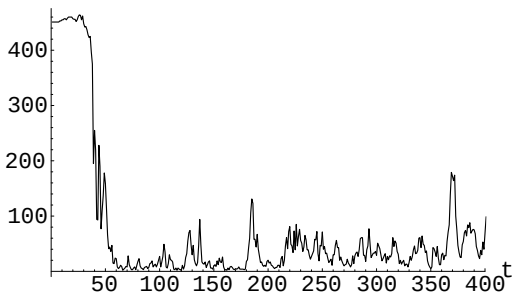
(d) $B_{\max}^{-1}$



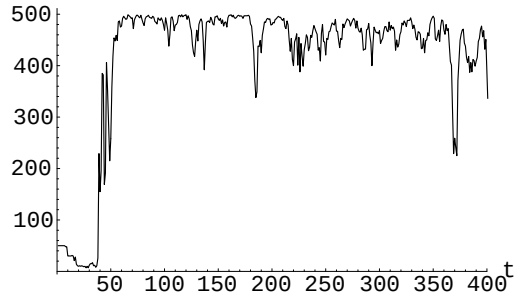
(e) $\max\left|\text{eigval}\,(B)\right|$



(f) $\max\left|\text{eigval}\,(B^{-1})\right|$



(g) Number of components



(h) # vertices in largest component

Figure 7.23: Iterator sample run: $n = 500$, $0 < t \leq 400$

```
{BestFitParameters → {amp → 41.9425, Aa → 0.174959, Bb → -0.0508046}, ParameterCITable →
          Estimate        Asymptotic SE      CI
 amp      41.9425         1.65465            {38.4662, 45.4188}
 Aa       0.174959        0.00567808         {0.16303, 0.186888}    ,
 Bb       -0.0508046      0.0670167          {-0.191601, 0.0899923}
EstimatedVariance → 16.197,
                                  DF      SumOfSq      MeanSq
            Model                 3       11941.7      3980.57
ANOVATable → Error                18      291.545      16.197 ,
            Uncorrected Total     21      12233.3
            Corrected Total       20      5749.4
                                    1.       0.357513  -0.330862
AsymptoticCorrelationMatrix → ( 0.357513      1.       -0.892126 ),
                                -0.330862  -0.892126     1.
                                                  Curvature
FitCurvatureTable → Max Intrinsic                 0.21883
                    Max Parameter-Effects         0.0901018 }
                    95. % Confidence Region       0.562552
```

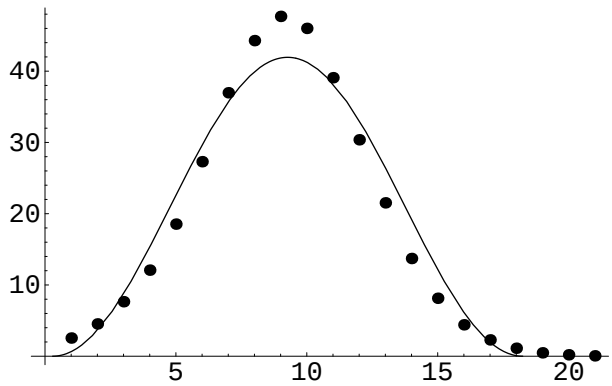
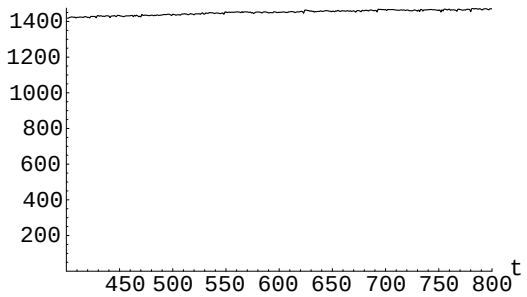
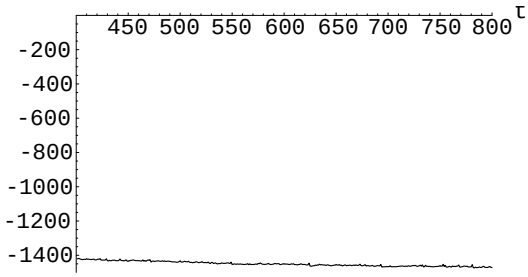


Figure 7.24: Case 4: Nonlinear best fit on B' at $t = 400$

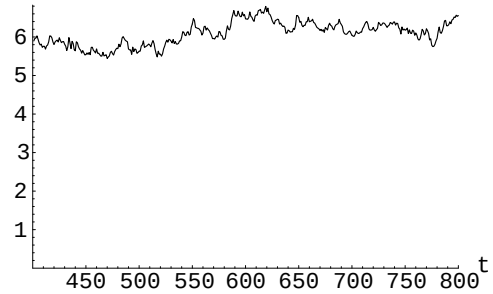
The giant component has 327 nodes, 369 edges, and 7 cycles with lengths $\{5, 7, 11, 15, 15, 42, 52\}$. Points are the respective D_k counts fitted to $f(x) = A \sin^2(ax + b)$ with $A = 41.9$, $a = 0.2$, and $b = -0.1$. Compare $D_{\max} = 48$ and $L = 21$ with the respective theoretical values of 47 and 19.



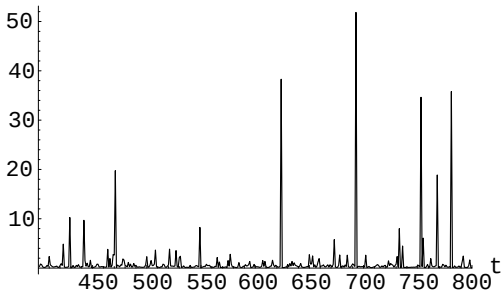
(a) $\ln\left(\det(B)\right)$



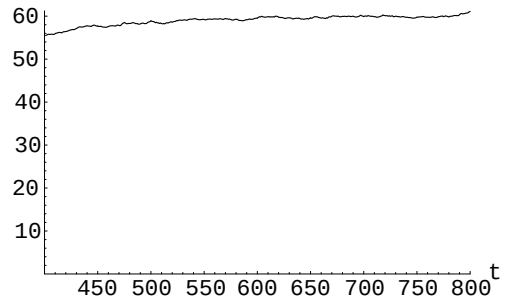
(b) $\ln\left(\det(B^{-1})\right)$



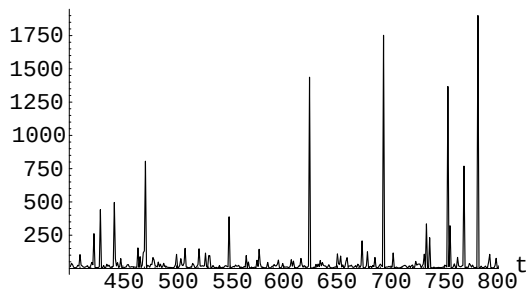
(c) $B_{\max}$



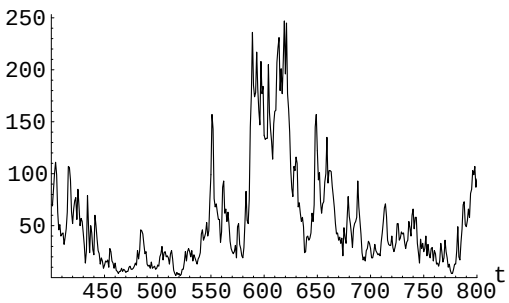
(d) $B^{-1}_{\max}$



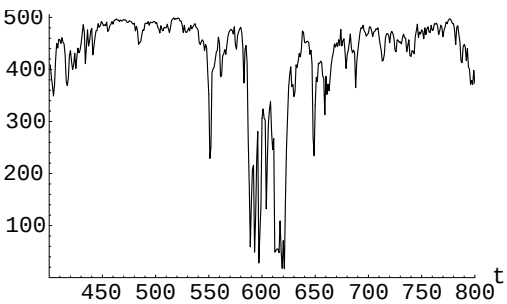
(e) $\max\left|\text{eigval}\,(B)\right|$



(f) $\max\left|\text{eigval}\,(B^{-1})\right|$



(g) Number of components



(h) # vertices in largest component

Figure 7.25: Iterator sample run: $n = 500$, $400 < t \leq 800$

```
{BestFitParameters → {amp → 43.0293, Aa → 0.17067, Bb → -0.19677}, ParameterCITable →

          Estimate      Asymptotic SE      CI
 amp      43.0293       1.3216             {40.2632, 45.7955}
 Aa       0.17067       0.00456054         {0.161124, 0.180215}    ,
 Bb       -0.19677      0.0565711          {-0.315175, -0.0783651}

 EstimatedVariance → 10.1747,

                                  DF      SumOfSq      MeanSq
              Model               3       12836.8      4278.92
 ANOVATable → Error               19      193.32       10.1747,
              Uncorrected Total   22      13030.1
              Corrected Total     21      5902.07

                                                  1.      0.399374   -0.3703
 AsymptoticCorrelationMatrix → ( 0.399374     1.      -0.910082 ),
                                  -0.3703   -0.910082     1.

                                                                  Curvature
 FitCurvatureTable → Max Intrinsic                                0.206014
                     Max Parameter-Effects                        0.110366 }
                     95. % Confidence Region                      0.565473
```

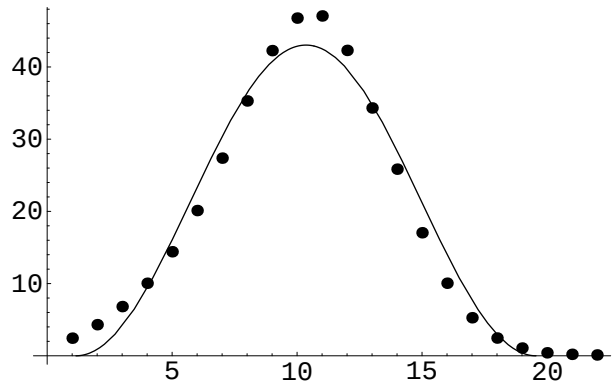
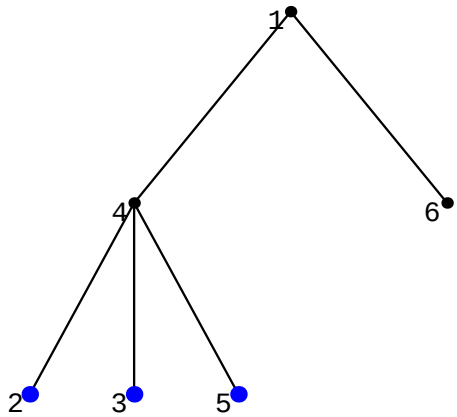


Figure 7.26: Case 4: Nonlinear best fit on B' at $t = 800$

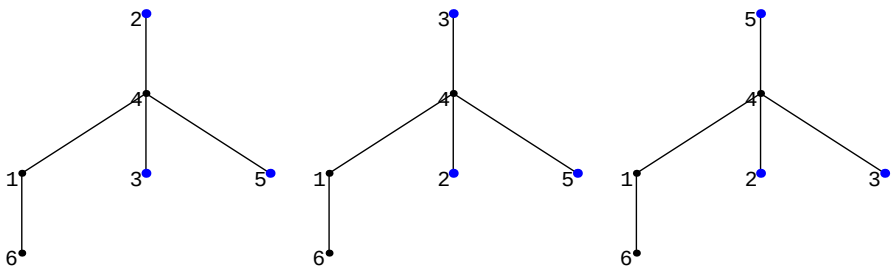
The giant component has 359 nodes, 396 edges, and 10 cycles with lengths $\{5, 6, 7, 7, 9, 11, 15, 17, 24, 44\}$. Points are the respective D_k counts fitted to $f(x) = A\sin^2(ax + b)$ with $A = 43.0$, $a = 0.2$, and $b = -0.2$. Compare $D_{\max} = 47$ and $L = 22$ with the respective theoretical values of 47 and 19.

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(a)



(b)



(c)

Figure 7.27: Singular values in tree-graphs
The adjacency matrix (a) is identical in rows/cols 2, 3, & 5 and thus singular (degenerate). Highlighted vertices in (b) emphasize the redundancy, while the three 'ranked embedding' views in (c) further illustrate their equivalence.

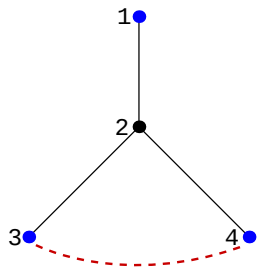


Figure 7.28: 4-tree graph
In the absence of the red-dashed link, the adjacency matrix of this tree is singular – highlighted vertices 1, 3, and 4 are identically situated. Introducing the additional link between nodes 3 and 4 breaks the kernel.

where δ gives the strength of the supplementary link. Notice, in particular, the locations of the inverse δ values within T^{-1}. The rôle of the binder in such a situation would take $T \to T - \alpha(T - T^{-1})$, giving

$$T_{\text{new}} = \begin{pmatrix} 0 & 1 & \frac{-\alpha}{\delta} & \frac{\alpha}{\delta} \\ -1 & 0 & 1-\alpha & 1-\alpha \\ \frac{\alpha}{\delta} & \alpha - 1 & 0 & \alpha\left(\frac{1}{\delta} - \delta\right) + \delta \\ \frac{-\alpha}{\delta} & \alpha - 1 & -\alpha\left(\frac{1}{\delta} - \delta\right) - \delta & 0 \end{pmatrix} \tag{7.3.3}$$

which places the $1/\delta$ values within the matrix proper. As $\delta \to 0$, relatively 'large' values (moderated by α) appear in the matrix corresponding to new strong links between the vertices associated with the impending null space, with such values rising very much faster than any drop in δ towards the singular state, as figure 7.29 on the next page shows, and the non-zero value of the determinant is preserved.

This provides *one* explanation for the peak values in the inverse maxima and principal eigenvalue plots on pages 244 and 246 (of course, equivalently, the least eigenvalue of the matrix proper simultaneously drops toward zero, but that behaviour is not exhibited so dramatically in plots) – they appear as a direct and immediate

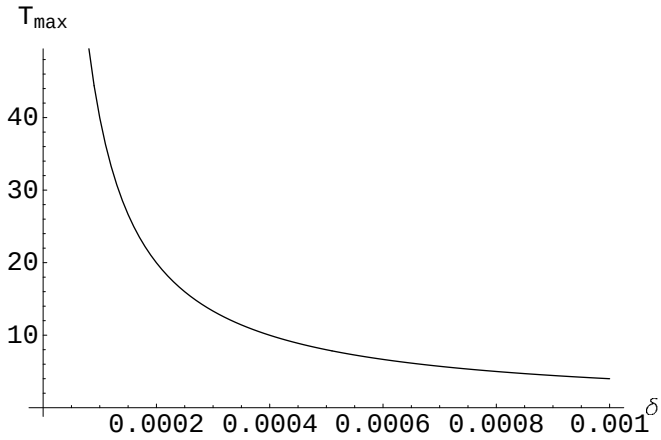
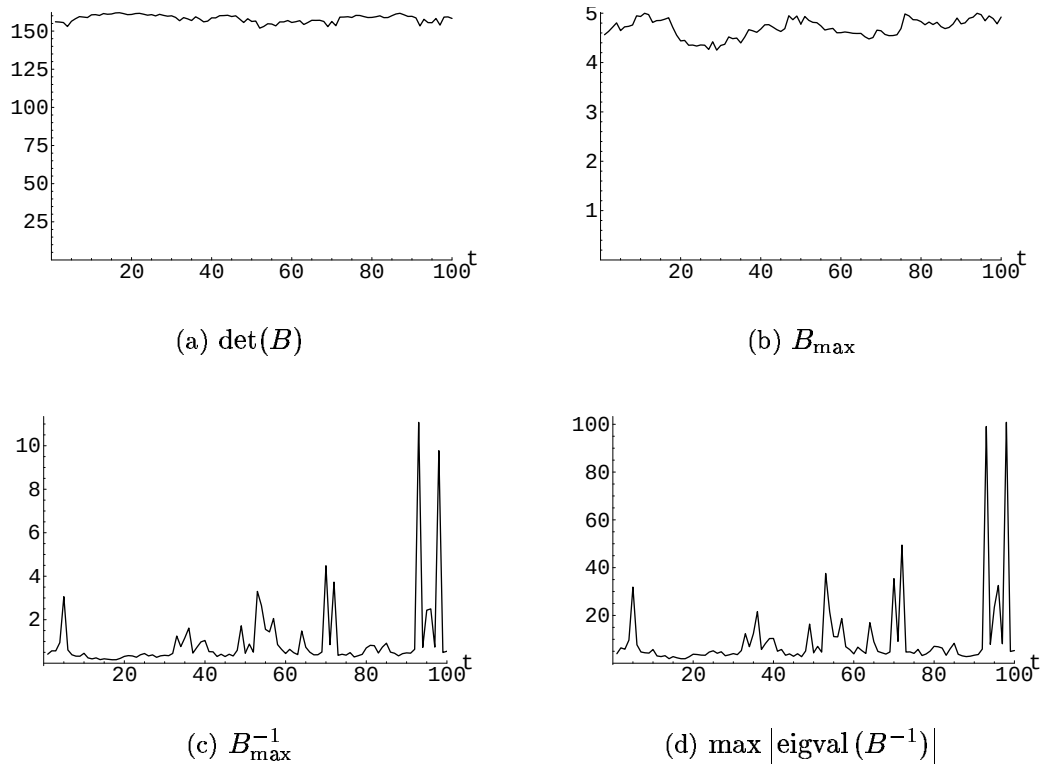


Figure 7.29: $T_{\max}$ versus δ
From equation 7.3.3, with $\alpha = 0.002$, the binder term in the HPS iterator equation reacts strongly to inhibit singular values in the matrix.

consequence of the binder's reaction to ill-conditioned, degenerate or singular states that may be occasioned by the stochastic SRN activity. The reason these peak values appear as spikes, rather than sustained quantities, is two-fold. Firstly, unlike the adjacency matrices of tree-graphs, or even augmented tree graphs, the HPS B-matrix as a whole has an extremely large determinant due to the presence of the SRN, which all but guarantees the freedom of the network from singular conditions, so that the matrix is well buffered from the effects of degenerate states within the embedded large-link structures. Secondly, the binder is highly effective in disallowing any tendency towards singular or ill-conditioned configurations – the swift 'kick' corrects the behaviour and eliminates the source of the disturbance, at the same time tending to strengthen the integrity of the emergent form in the network structure.

To reinforce and affirm the explanation, a small HPS model ($n = 80$) was allowed to run through its start-up phase and then the B-matrix values were tracked through a number of iterations until a short sequence containing a target region free of large

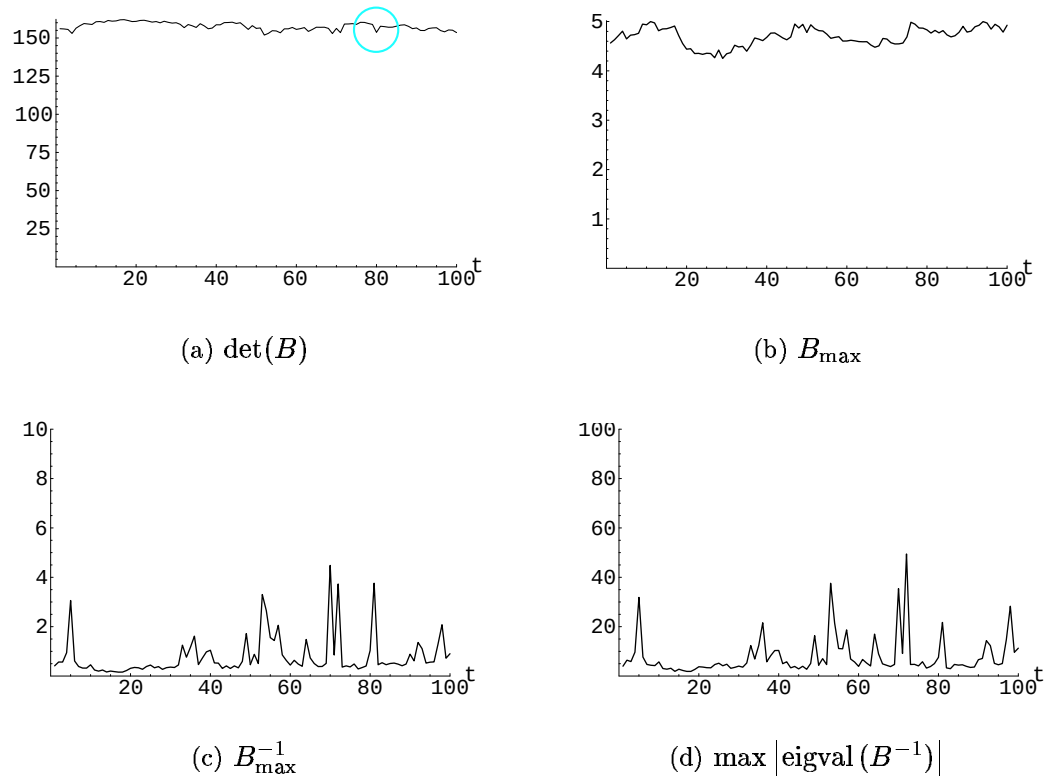
B^{-1} values was observed. The B-matrix at the start of that sequence was captured, as were the attendant SRN values for the duration of the sequence. A singular state was induced in one B within the target region by pruning a leaf node and then introducing a δ-value to reconnect that node with the large component so as to control the extent to which the singular states were restored and thus control the impact on the kernel. Finally, the sequence was replayed with the modifications in place and for various values of delta. The results are shown in figures 7.30 to 7.35.



(a) $\det(B)$ (b) $B_{\max}$

(c) $B^{-1}_{\max}$ (d) $\max\left|\text{eigval}\left(B^{-1}\right)\right|$

Figure 7.30: Original sequence without intervention
Note the interval $75 < t < 95$, which is relatively free of large B^{-1} values: this is the target region for subsequent intervention.

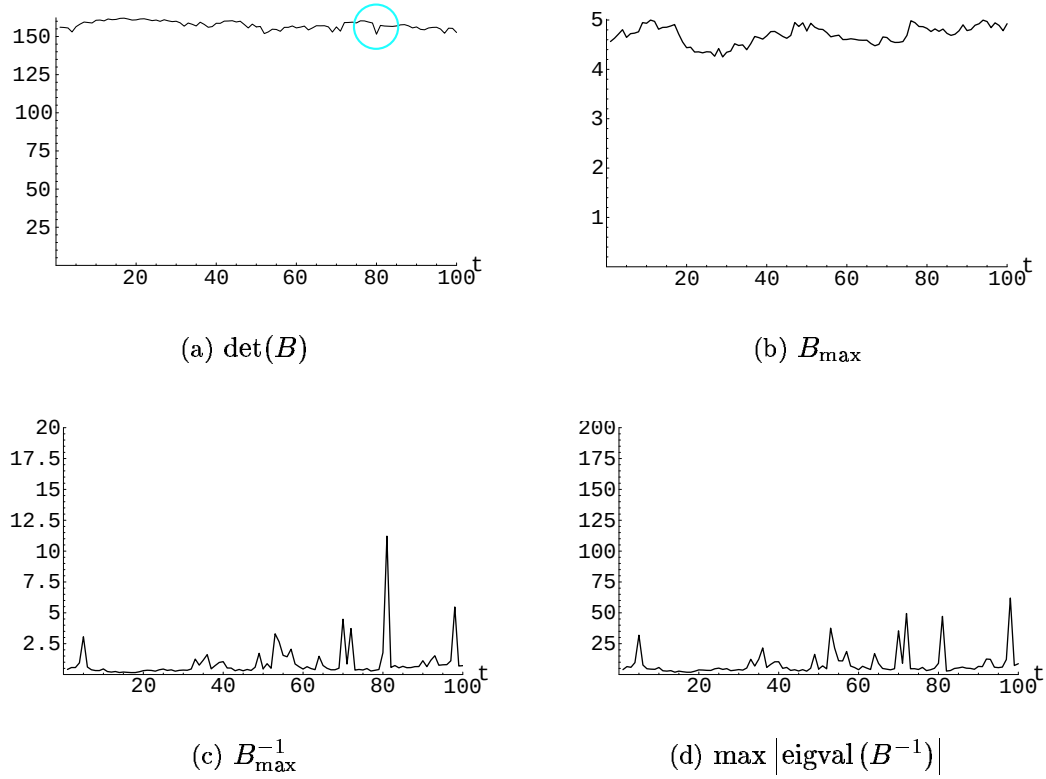


(a) $\det(B)$ (b) $B_{\max}$

(c) $B_{\max}^{-1}$ (d) $\max\left|\text{eigval}\left(B^{-1}\right)\right|$

Figure 7.31: Replayed sequence with intervention: $\delta = 10^{0.5}$ at $t = 80$
In (a) the determinant shows a small dip (circled), indicating the occurrence of the intervention, and the spikes in $B_{\max}^{-1}$ and the principal eigenvalue of B^{-1} at $t = 81$ demonstrate the binder's reaction.

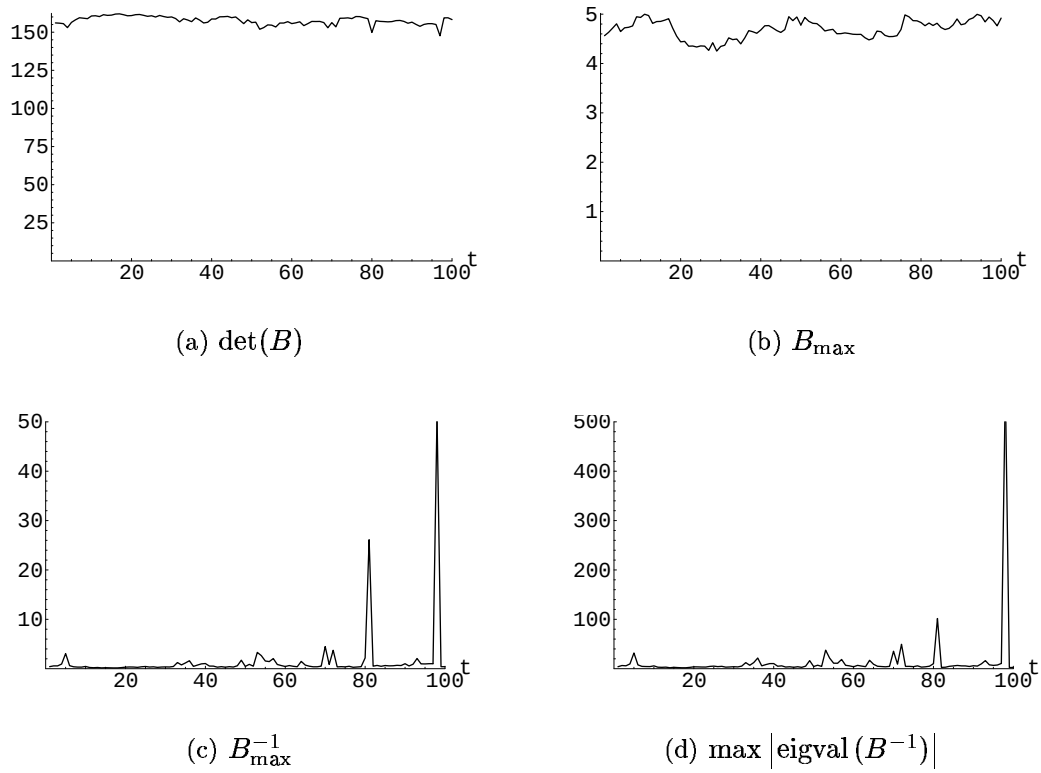
Note, also, that while the sequence is identical to the original up to $t = 80$, the single introduced change to the history alters the 'timeline' of the network: the large spikes in the original sequence in the interval $95 < t < 100$ have dropped markedly, due solely to this one event (the SRN values being identical throughout). This is reflected, also, in $\det(B)$ for $t > 80$.



(a) $\det(B)$ (b) $B_{\max}$

(c) $B^{-1}_{\max}$ (d) $\max\left|\text{eigval}\left(B^{-1}\right)\right|$

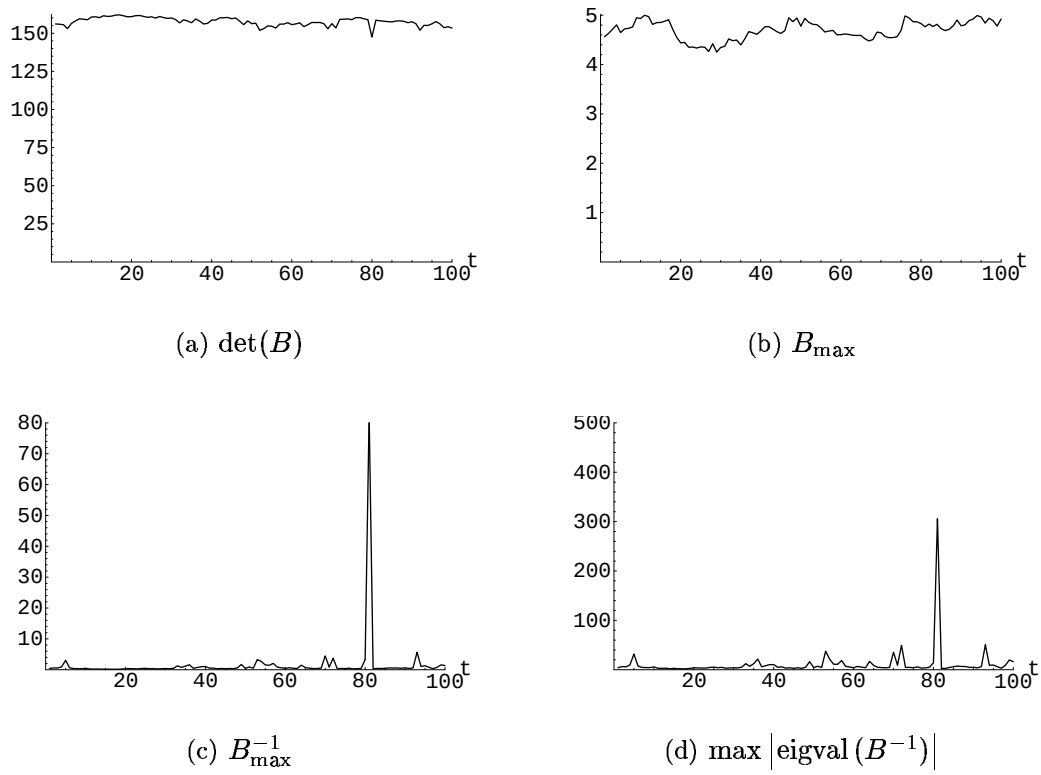
Figure 7.32: Replayed sequence with intervention: $\delta = 10^{0.25}$ at $t = 80$
The momentary dip in the determinant, again circled in (a), is more marked than previously and the induced peaks in (c) and (d) are correspondingly greater, reflecting the smaller δ-value. The later ($t > 90$) activity is stronger, also.



(a) $\det(B)$ (b) $B_{\max}$

(c) $B^{-1}_{\max}$ (d) $\max\left|\text{eigval}\left(B^{-1}\right)\right|$

Figure 7.33: Replayed sequence with intervention: $\delta = 10^{0.125}$ at $t = 80$
As δ is taken smaller, the effects on the network via the reaction of the binder become increasingly pronounced. A particularly large peak in the B^{-1} data is observed at $t \sim 100$, showing that events ripple forward in time, interacting with the SRN with unpredictable consequences and illustrating the extent to which the time development of the system is influenced by prior phenomenology.



(a) $\det(B)$ (b) $B_{\max}$

(c) $B^{-1}_{\max}$ (d) $\max\left|\text{eigval}\left(B^{-1}\right)\right|$

Figure 7.34: Replayed sequence with intervention: $\delta = 1$ at $t = 80$

The effect on the determinant is now quite pronounced and the consequent B^{-1} values are very strong, though the unity value of δ is well above that required to establish a true null space and not yet small enough to provoke maxima in B greater than extant values, as evidenced by (b) in this and the preceding figures. This suggests that the matrix is being reinforced via the binder reaction but not to such extent that the general characteristics of the network are impugned.

Plotting the principal eigenvalues of the B^{-1} matrix against values of delta (figure 7.35) provides results consistent with those obtained analytically for the 4×4 matrix given in equation 7.3.3 on page 249.

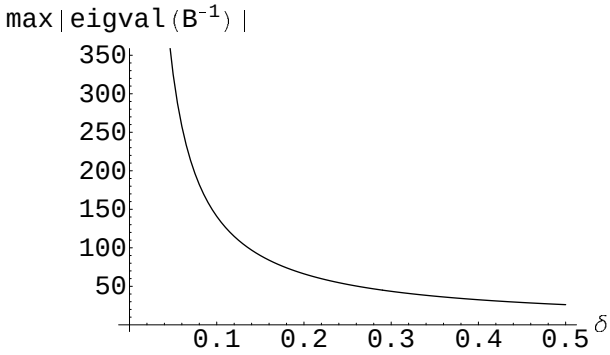


Figure 7.35: B^{-1} sensitivity to δ-induced singular states
consistent with the theoretical result shown in figure 7.29 on page 250, actual results of intervention in B show the sensitivity of B^{-1} in the binder term to values of δ, the parameter used to induce and control singular states.

However, when comparing this plot with figure 7.29 on page 250, which shows the effect of changing δ on the matrix maxima, the lack of sensitivity in the $B_{\max}$ values to changes in δ (illustrated in the preceding figures) seems anomalous. The matter is resolved by recognizing that the earlier result obtained from a very small matrix and with no SRN term in the recursion. In the later case, even though the system is still relatively small, at $n = 80$ nodes, there is sufficient activity elsewhere in the matrix, largely unaffected by the induced disturbance from the δ influence, which, together with the presence of SRN, provides a noisy background and competing local maxima to buffer the system from isolated deviations.

A *second* explanation for the intermittent peaks in the B^{-1} data is now apparent. Not only are there immediate responses from the binder to singular or ill-conditioned states but such reactions, in conjunction with the stochastic values, may have a

flow-on effect – in a sense propagating along the forward 'time-line', as displayed in figures 7.31 to 7.34 – resulting in the subsequent appearance of further peaks in the principal eigenvalues and maxima of the matrix inverse which, in turn, may trigger yet more events. In this manner, then, the binder serves not merely as a regulator and restorative force in the network but also as an active and creative generator of phenomena.

Having demonstrated a mechanism, beyond that of the SRN alone, for the generation of new links in the context of a B-matrix with a solitary large component, it becomes pertinent to examine how this mechanism responds to the presence of two or more components of non-trivial size.

The effect of the binder on multiple components and the process of network construction

As a prelude to investigating how the binder mechanism responds to multiple large components, a simpler case was first examined: that which obtains from presenting to the binder a graph comprising a single random tree plus a number of disassociated nodes. Following the previous discussion, the presence of disassociated nodes is sufficient, in the absence of a stochastic background, to render the graph singular, regardless of the condition of the connected component. In view of the binder's immediate and vigorous reaction to counter singular tendencies, it should be expected that new edges will be generated in response to the presence of the extra nodes. What is not clear, from that expectation alone, is whether the binder discriminates to give preferential treatment to extend the tree size by incorporating the spare nodes (while leaving that extended component as a tree sub-graph) or whether singular values in

the tree component take priority. Or, indeed, whether there is any distinction between the two cases.

The following method was employed:

1. generate a random tree graph and extract its $n \times n$ adjacency matrix to use as a 'seed' in a larger matrix

2. construct an empty $m \times m$, $(m > n)$ matrix and 'plant' the seed tree within it; plot the corresponding graph and also compute the kernel of the adjacency matrix;

3. antisymmetrize the resulting matrix to comply with HPS conditions

4. generate an $m \times m$ antisymmetric random matrix (the SRN term) and construct the matrix sum of the two matrices

5. compute the inverse of the resultant matrix

6. pass the inverse through a threshold filter to generate a sparse $m \times m$ adjacency matrix whose entries are zero except for unit-valued elements corresponding to elements of the inverse that exceeded the threshold value; record the number of edges and their associated vertices

7. construct the matrix sum of this sparse adjacency matrix and the original $m \times m$ adjacency matrix

8. plot the corresponding new graph and compare with the original to visually identify affected nodes and new links

9. compare the list of affected nodes with the nullspace (kernel) of the original structure

10. repeat step 6 for a range of threshold values

In effect, this procedure demonstrates the operation of $(B + B^{-1})$ in isolation, with no new SRN term to obscure its behaviour, over a single iteration. The following results are typical; here the seed tree contains 40 nodes and there are 10 additional disassociated nodes. Figure 7.36 on the current page shows the original graph, while figure 7.37 shows the modified graph after action by the binder mechanism. Every affected node, highlighted in blue, is associated with the kernel of the original graph and the binder has acted on those nodes indiscriminately, in the sense that no obvious preference is given to vertices regardless of their situation.

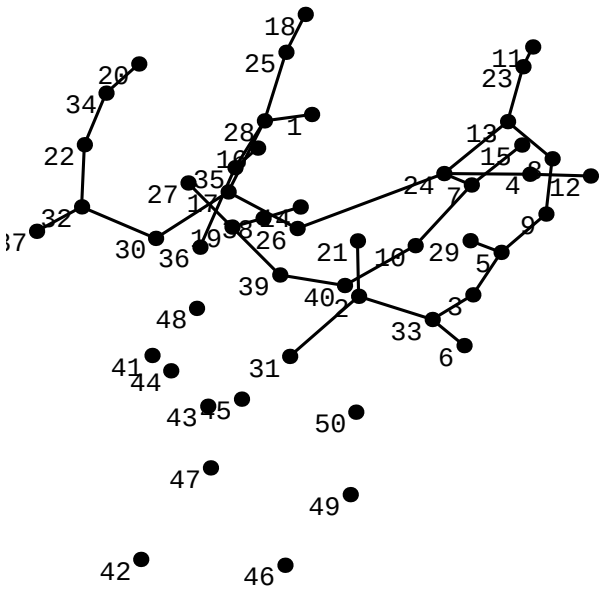


Figure 7.36: Random tree graph ($n = 40$) plus 10 disassociated nodes
The kernel of this graph is associated with nodes {3, 6, 9, 10, 11, 13, 15, 16, 20, 21, 22, 26, 27, 29, 30, 31, 36, 37, 39, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50}

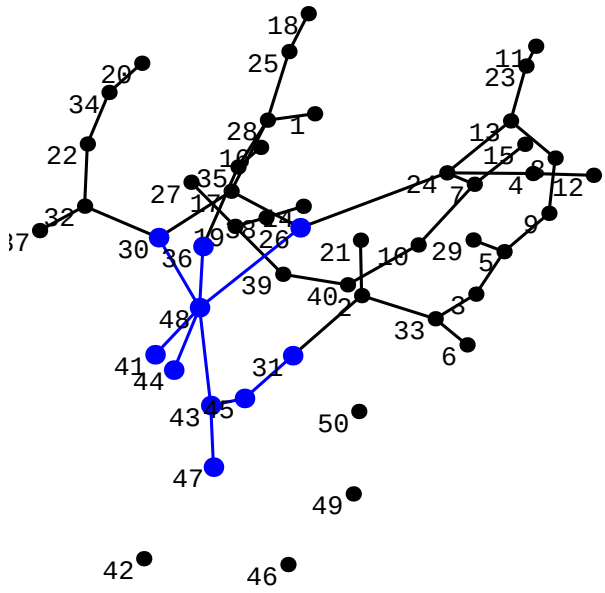


Figure 7.37: Binder-modified graph
Nodes {26, 30, 31, 36, 41, 43, 44, 45, 47, 48} and the 9 edges highlighted in blue have been acted upon by the binder using a large-link threshold of 70%.

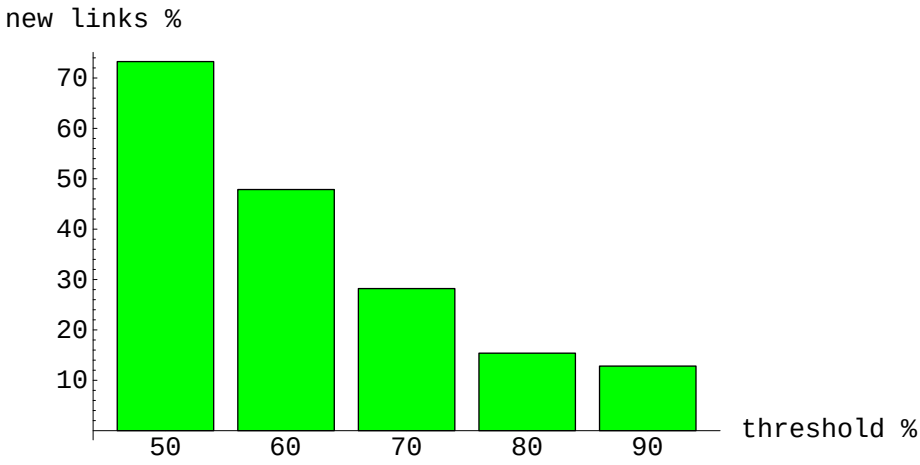


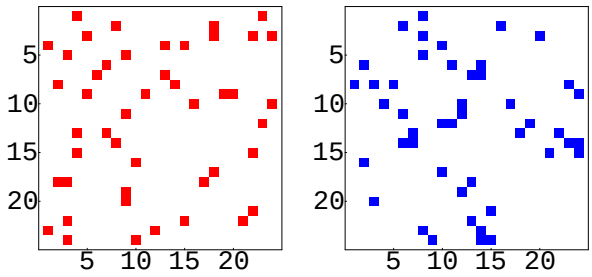
Figure 7.38: Histogram: binder-generated edges versus large-link threshold
The number of new links is given as a percentage ratio of edges in the original graph. Category values for *threshold* are centred in the range.

Figure 7.38 provides a histogram showing the percentage of additional links generated by the binder for threshold values ranging from 45% to 95% of the maximum link magnitude. Inspection of the raw data shows that for threshold values of about 50% and above, the *only* vertices affected are those that are associated with the null space or kernel of the original graph. Cross-linking of vertices within the tree component is not observed until the threshold is made sufficiently small, where 'sufficiently small' is to be interpreted as meaning that most, if not all, former kernel nodes have been excised from the null space.

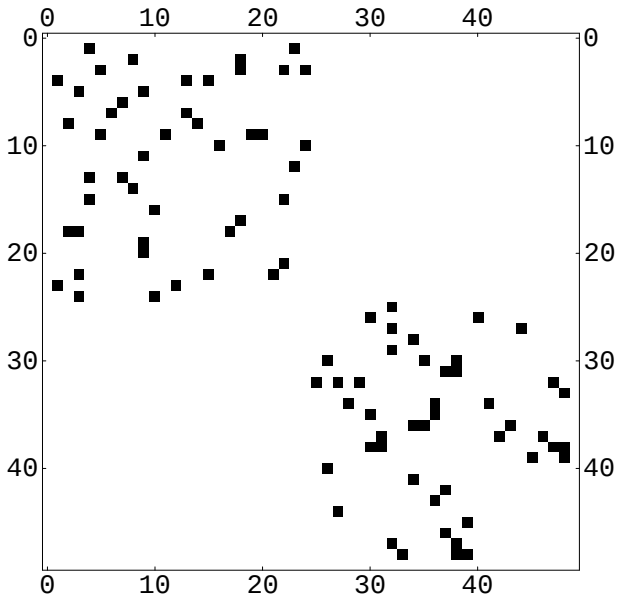
While this is strong behaviour, its impact within the iterator as a whole is tempered by the binder strength parameter, α. Nonetheless, this attribute of the binder demonstrates that within an HPS there exist nodes that receive preferential treatment in attracting new links – one might say that such nodes are '*sticky*' [361].

Proceeding now to examine the case where a matrix comprising two large components is presented to the binder, essentially the same method was again employed but with two random tree subgraphs, rather than one, planted within a larger empty matrix. Figure 7.39 on the following page shows the initial set up for a typical trial and figure 7.40 on page 263 shows spring-embedding representations of the graph before and after the intervention of the binder. As with the preceding case, the binder targets kernel nodes in the original graph and preferentially introduces new links to those nodes with the extent of that activity governed by the large-link threshold, which was set at 75% in this example.

Figure 7.41 on page 264 illustrates the location of new links in the previously vacant off-block-diagonal regions of the adjacency matrix, indicating that the two original components are no longer disconnected but form a single giant component.



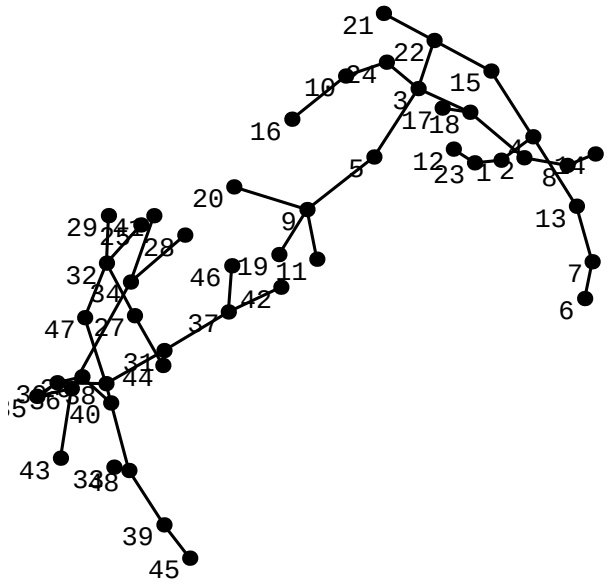
(a) 'Seed' components



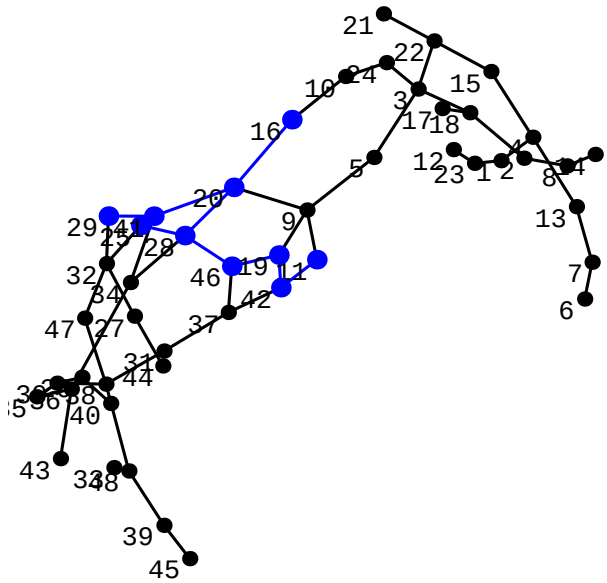
(b) Full adjacency matrix

Figure 7.39: Combining two (24×24)-node random tree subgraphs
These are density plot representations of the respective adjacency matrices for the two 'seed' trees (a) and the corresponding representation (b) of the total (48×48) matrix incorporating the seed components. Note the block diagonal form of the latter.

The kernel of the first component involves vertices {1, 2, 5, 6, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 24} while the null space of seed 2 involves vertices {25, 28, 29, 31, 41, 42, 46, 47}. The set of kernel nodes for the entire graph is the union of these two sets.



(a) Graph with two disconnected components



(b) Binder-modified graph

Figure 7.40: Effect of the binder on two (24 × 24)-node components
The kernel of (a) is associated with nodes {1, 2, 5, 6, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 24, 25, 28, 29, 31, 41, 42, 46, 47}. In (b), highlighted vertices and links show where the binder has been active, introducing 10 new edges over nodes {11, 16, 19, 20, 25, 28, 29, 41, 42, 46}, all of which belonged to the kernel of the original full graph shown at (a). The large link threshold here was set at 75%.

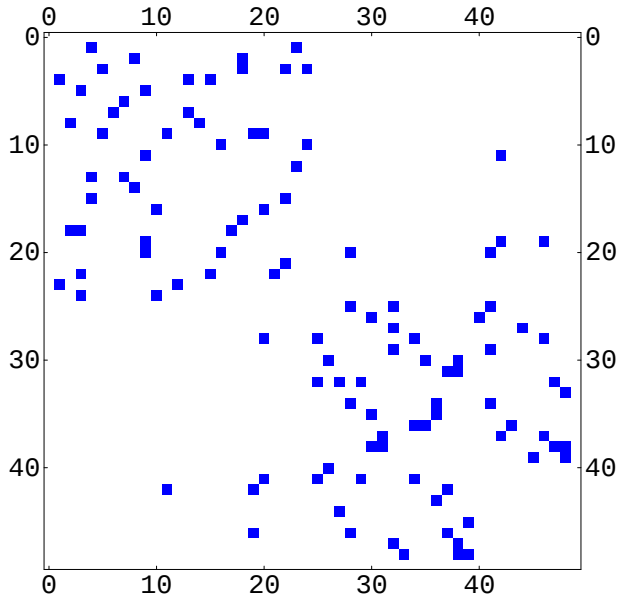
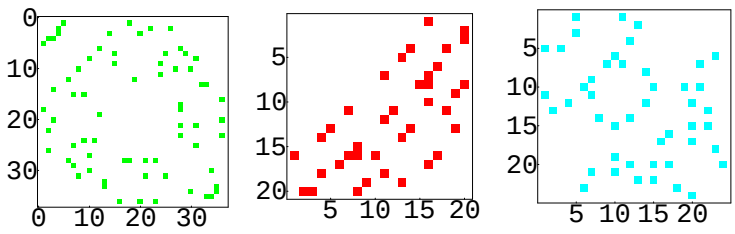


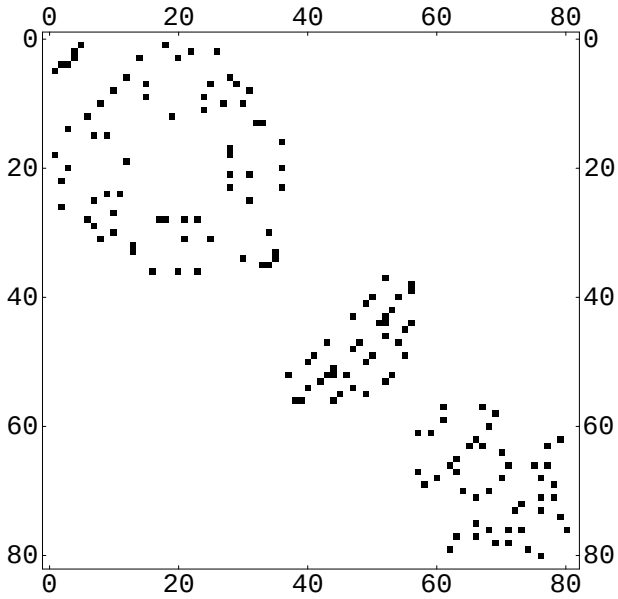
Figure 7.41: Case: 2-components – density plot of the final adjacency matrix Compare this with figure 7.39(b) to see the location of new links in the previously vacant off-block-diagonal regions of the adjacency matrix.

After just one iteration, however, this is only a very weakly augmented tree graph but, on the strength of results reported earlier, one would expect further iterations with the SRN at work to complete the task of providing connectivity consistent with the S^3 form.

Next, the procedure was extended to examine a graph with three block-diagonal components, tree subgraphs all, comprising 36, 20, and 24 nodes respectively. The results shown in figures 7.42–7.44, on pages 265–267, are consistent with previous observations and, again, the binder is seen to unite previously disconnected components, acting only on nodes in the null space of the original graph. This is not always the case, as illustrated by a further example from the same procedure, shown in figure 7.45 on page 269. Blue highlights in 7.45(b) demonstrate the binder acting to connect two of the three components via 9 new links affecting 7 vertices, leaving



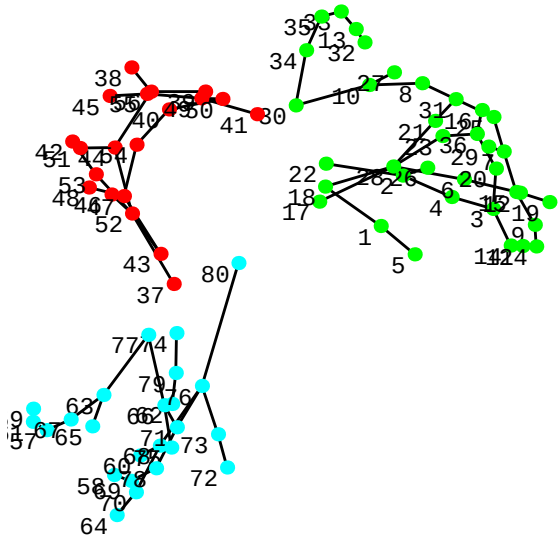
(a) 'Seed' components



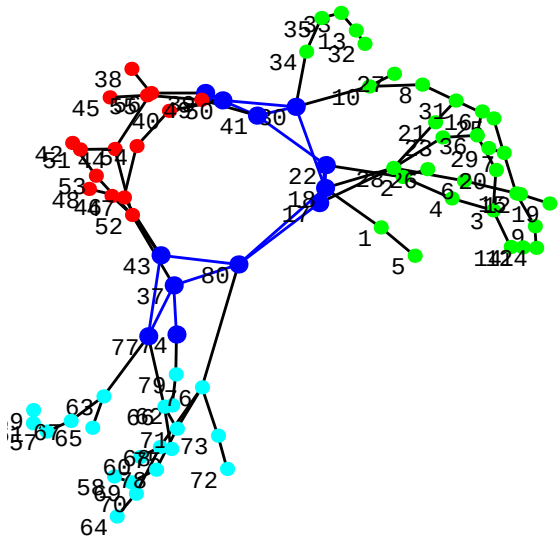
(b) Full adjacency matrix

Figure 7.42: Combining three random tree subgraphs
(a) shows density plot representations of the respective adjacency matrices for the three 'seed' trees and the corresponding representation (b) of the total (80 × 80) block-diagonal matrix incorporating the seed components.

The kernel of the first component involves vertices {4, 5, 6, 8, 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 29}, the null space of seed 2 involves vertices { 37, 38, 39, 41, 43, 46, 48, 50, 54}, and the null space of seed 3 involves vertices {62, 65, 74, 75, 77}. The set of kernel nodes for the entire graph is the union of these sets.



(a) Three disconnected components



(b) Binder-modified graph

Figure 7.43: Effect of the binder on three disconnected components (i) The kernel of (a) is associated with nodes {4, 5, 6, 8, 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 29, 37, 38, 39, 41, 43, 46, 48, 50, 54, 62, 65, 74, 75, 77}. Blue vertices and links show the effect of the binder (threshold 75%) – 15 new links over 10 nodes {17, 18, 22, 30, 37, 39, 41, 43, 50, 74, 77, 80}, all belonging to the original kernel. This set is identical to the complement of the kernel of the matrix inverse.

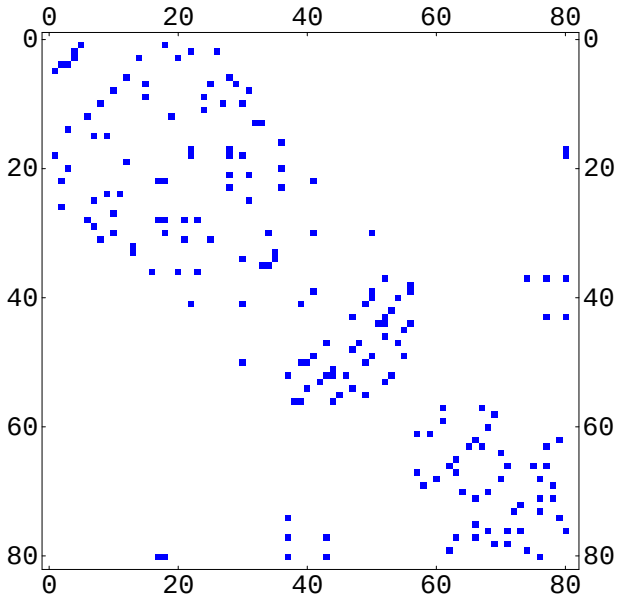


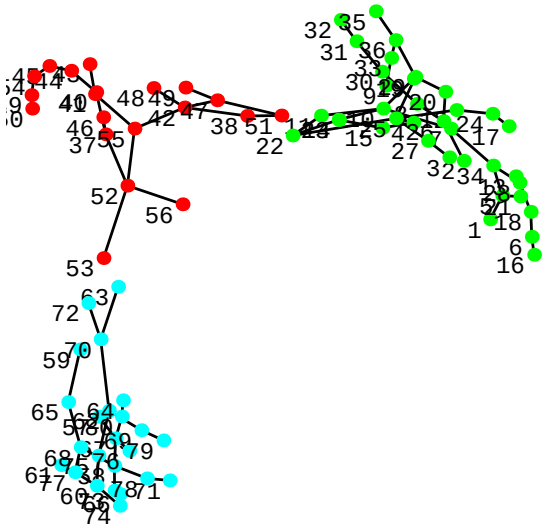
Figure 7.44: Case: 3 components – density plot of the final adjacency matrix showing the location of new links in the previously vacant off-block-diagonal regions.

the third component disconnected. Analysis of the component adjacency matrices reveals that the null space of the neglected subgraph involves just 4 of 24 vertices, whereas the kernels of the other two components involve 15 of 36 nodes and 8 of 20 nodes respectively. It may be inferred, therefore, that the lone component contributed least to the degeneracy of the whole graph and so escaped (for the time being) the efforts of the binder to resist singular states. After the first step, the null space of the whole graph has reduced in size from 27 to 18 nodes. A second pass through the binder unifies the graph and the highlights in figure 7.45(c) identify the binder activity that captures the third component via a further 4 links involving 5 nodes, which includes a new cross-link between nodes 59 and 63. Notice that node 59 does not appear in the kernel and this is the only instance, in this example, of binder activity affecting a non-kernel vertex. After the second step, the null space

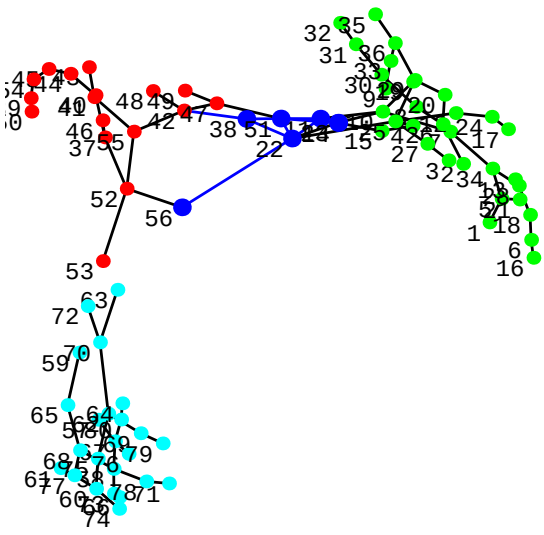
has increased in size, now involving 31 nodes, {1, 2, 4, 11, 13, 14, 16, 18, 19, 20, 21, 22, 23, 38, 43, 48, 49, 51, 53, 55, 57, 58, 63, 65, 66, 70, 71, 72, 74, 75, 76}. Seven of these ({58, 65, 66, 70, 71, 74, 75}) were not previously associated with singular states, indicating that the change in connectivity has caused these nodes to acquire a configurative similarity with other nodes that was previously lacking. This is to be interpreted as indicating that the unified graph remains very active, or *volatile*, in the sense that it will attract more new links when next exposed to the action of the binder. That is, if this network was but a part of a larger system then, in the absence of similarly conditioned components, it would likely acquire new cross-links, further augmenting the tree structure and taking the graph closer to the expected S^3 form. On the other hand, if other volatile components were present, links with those components would be highly likely within just a few iterations, as would the appearance of cross links, with the behaviour occurring selectively according to the extent to which the components or their constituent nodes take the whole network towards a singular state.

7.3.3 Distribution of active ('sticky') nodes

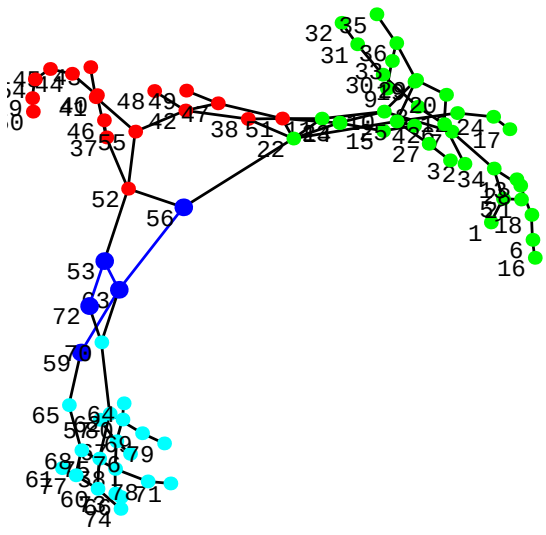
Having ascertained that vertices associated with singular states or degenerate tendencies – that is, nodes involved in the actual or latent null space or kernel – attract the attention of the binder mechanism, thus appearing more active in the network than other vertices, it became germane to examine their distribution in terms of the distance measure defined on the HPS. At each of 100 iterations of a 150-node HPS that had first been allowed to reach its equilibrium plateau, the pair-wise shortest paths between active nodes in the largest connected component of B' (obtained from



(a) Three disconnected components



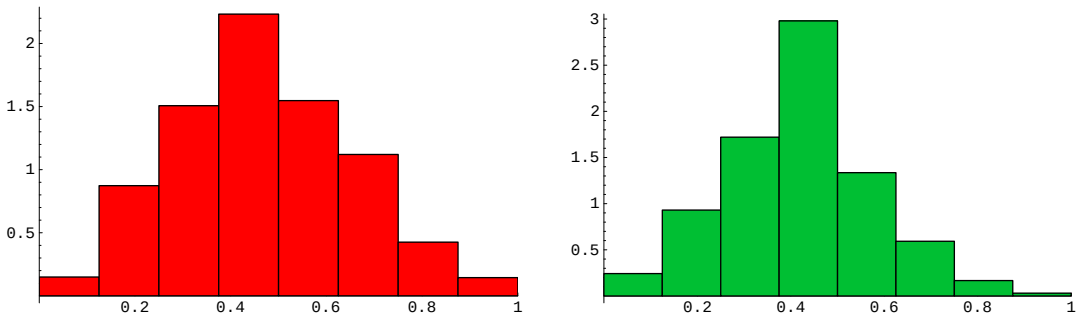
(b) Binder-modified graph - step 1



(c) Binder-modified graph – step 2

Figure 7.45: Effect of the binder on three disconnected components (ii) The kernel of (a) is associated with nodes {1, 2, 4, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 38, 43, 48, 49, 51, 53, 55, 56, 57, 63, 72, 76}. Blue vertices and links in (b) and (c) show the effect of the binder (threshold 75%) in two passes. After the first step (b), one component remains disconnected and the kernel has reduced to nodes {1, 2, 11, 13, 16, 18, 19, 20, 21, 38, 43, 48, 53, 55, 57, 63, 72, 76}. The second step (c) incorporates the third component.

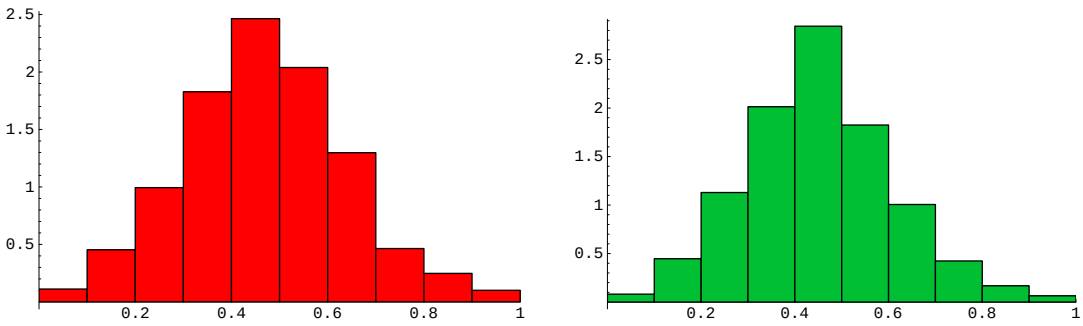
coarse-graining the HPS matrix as before) were computed. Concurrently, so too were the pair-wise shortest paths between *all* nodes. The respective path lengths were recorded, after normalization with respect to the longest path in each instance, and counted in aggregate to produce the histograms in figures 7.46 ($n = 150$) and 7.47 ($n = 400$).



(a) Active nodes (b) All nodes

Figure 7.46: Histograms: normalized vertex distance distributions (i) with normalized counts on the vertical axis against normalized pair-wise 'distances' on the horizontal. (a) shows the distribution of distances between active ('sticky') nodes, while (b) gives the distribution of distances between all nodes. Data were collected from a HPS with $n = 150$ over 100 iterations after the system had attained its equilibrium plateau.

In each of these cases, the pair-wise distances are approximately normally distributed (the approximation being stronger in the second figure, where the size of the underlying B-matrix is substantially greater) and, in particular, the distribution of pair-wise distances between active nodes (a) is not significantly different from the distribution between all nodes (b). That is, it may be inferred that the new links created by the binder mechanism (whether serving to internally cross-link components or to extend and connect components) are, by virtue of their most frequent association with active nodes, essentially randomly distributed across the network. This is a highly



(a) Active nodes (b) All nodes

Figure 7.47: Histograms: normalized vertex distance distributions (ii) for comparison with the preceding figure 7.46. Data were obtained by the same procedure and under the same conditions except that here $n = 400$.

significant feature, for reasons that will be revealed in due course.

7.3.4 Embeddability testing – S^3 in E^4

Previous tests for emergent intrinsic geometrical structure, in both HQS and HPS, have involved the examination of the depth structure of the graphs obtained by coarse-graining samples of the B-matrices and, of course, the ideas on 'embeddability' put forward in §6.8.5 are as relevant to HPS theory as they were to HQS theory. That is, the theoretical analysis provided the expectation that the simplest and most probable global structure that arises should be consistent with the geometry of the hypersphere S^3, in that the dominant links define a structure that is strongly embeddable in S^3 whereas weaker links tend to diffuse that embedding, with the hypothesis that gebits have a 'fuzzy' embeddability in S^3.

While the results of earlier tests were suggestive, the interplay between 'strong' and 'weak' links is too subtle for the fuzzy embeddability hypothesis to be tested

adequately via the coarse-graining procedure. To test this hypothesis by other means required a method to allow the link variables to determine an ordering of a gebit's constituent monads in terms of a known yet independent mathematical construct and thereby reveal any geometrical or topological properties inherent to the gebit.

Consider: the surface of a $(m-1)$-sphere of radius r in m-dimensional Euclidean space defines a locus of points satisfying the generalized Pythagorean theorem

$$\sum\nolimits_{\alpha=1}^{m} \left(x_{\alpha}^{(i)} - x_{\alpha}^{(0)}\right)^2 - r^2 = 0, \tag{7.3.4}$$

where $X^i = \left\{x_1^{(i)}, \ldots, x_n^{(i)}\right\}$ defines the coordinates of the i^{th} point and X^0 is the centre of the $(m-1)$-sphere. Hence, just as an ordinary 2-sphere[3] is precisely embeddable in three dimensions (illustrated by figure 7.48 on this page), a 3-sphere is precisely embeddable in four (or more) dimensions.

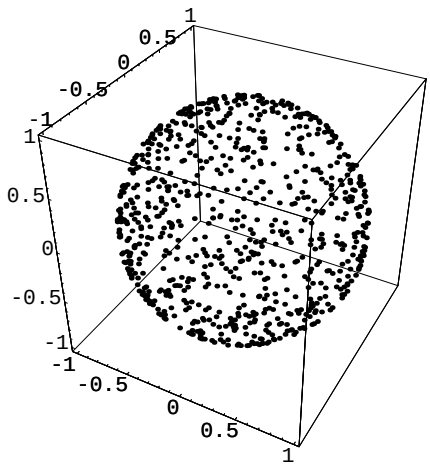


Figure 7.48: Representation of S^2 in E^3
illustrating the embedding of points with a near-S^2 signature in a higher dimensional Euclidean space (in this case, E^3).

One may consider the link variables B_{ij} as defining an n-dimensional network

[3]Here, the terminology of topologists is adopted – geometers, following Coxeter (generally regarded as the greatest geometer of the 20th Century), call an ordinary sphere a '3-sphere'.

of 'springs' between i, j pairs of monads, so that the spring between monad i and monad j has spring constant $\kappa = B_{ij}^2$ and natural length $1/|B_{ij}|$. Then, with X^i as the coordinates of monad i in the m-dimensional Euclidean space E^m, let

$$d_{ij} \equiv d(X^i, X^j) = \sqrt{\sum_{\alpha=1}^{m} \left(x_\alpha^{(i)} - x_\alpha^{(j)} \right)^2} \tag{7.3.5}$$

be the Euclidean distance between monad i and monad j in that E^m and define

$$V(X) = \sum_{i>j} B_{ij}^2 \left(d_{ij} - \frac{1}{|B_{ij}|} \right)^2 \tag{7.3.6}$$

so that $V(X)$ is an embedding measure corresponding to a '*spring embedding*' model such that the measure is zero if and only if (since singular B_{ij} are not permitted by the iterator) the coordinates of each monad provide separations that correspond precisely to the respective natural spring lengths.

An arbitrary assignment of coordinates stores 'energy' in the network and $V(X)$ is minimized by a procedure that lowers the energy by attempting to place strongly linked monads at short separations, whilst allowing greater separations for weaker links (represented by weaker and longer springs that have less influence and which, in effect, diffuse the E^m embedding). As a result, this provides a means of embedding gebits in Euclidean spaces (being known yet independent mathematical constructs) with no constraints other than those imposed by the information carried by the link variables. The ensuing mapping of coordinates may be searched for the hypothesized hypersphere signature.

The following method was used for this analysis:

1. choose a dimension for the embedding ($m = 2$, 3, or 4)

2. initialize a $n \times n$ HPS B matrix and apply the iterator procedure through the start-up phase

3. after the start-up phase, to each of the n monads, randomly assign Euclidean coordinates x_k, $(k = 1 \ldots m)$

4. compute the 'centre of mass' (cm) position (i.e. the origin in the Euclidean space) of the network defined by $x_\alpha^{\mathrm{cm}} = (1/n) \sum_{i=1}^{n} x_\alpha^{(i)}$

5. compute the pair-wise Euclidean distances between all monads

6. compute the total spring 'force' on each monad and the corresponding component force in each coordinate by

$$F_{x_\alpha}^{(i)} = \sum_{j \neq i} \left[B_{ij}^2 \left(d_{ij} - \frac{1}{|B_{ij}|} \right) \frac{\left(x_\alpha^{(i)} - x_\alpha^{(j)} \right)}{d_{ij}} \right] \tag{7.3.7}$$

7. iteratively, compute the total energy in the network (using equation 7.3.6) and perturb the network by applying micro-adjustments to the coordinates of each monad so that

$$x_\alpha^{(i)} \rightarrow x_\alpha^{(i)} - \frac{a.F_\alpha^{(i)}}{F_{\mathrm{M}ax}} \tag{7.3.8}$$

(where a is a control parameter such that $a/F_{\mathrm{M}ax}$ dynamically scales the repositioning of the monad) while the energy gradient is negative, indicating reduction in total energy; stop when the sign of the gradient changes, indicating the existence of a (local) minimum

8. compute the distance of each monad from x_α^{cm} by

$$r^{(i)} = \sqrt{\sum_{\alpha=1}^{m} \left(x_\alpha^{(i)} - x_\alpha^{\mathrm{cm}} \right)^2} \tag{7.3.9}$$

9. extract the radial distribution of the monads in E^m

10. analyze the distribution of radial distances, looking for evidence of a preferred or most likely, radius – that is, evidence of a S^{m-1} signature

11. repeat the analysis for randomly assigned coordinates *without* perturbations and the energy minimization; compare with previous results

Minimization procedure trials

To test the minimization procedure and to illustrate 'ideal' signatures, trial B-matrices were constructed such that their entries satisfied $\left(d_{ij} - \frac{1}{|B_{ij}|}\right) = 0$ when the d_{ij} obtained from assigning to each monad coordinates corresponding to a unique position on an exact $(m-1)$-sphere for the cases $m = 2, 3, 4$. In each trial, the coordinate set was then randomized and, with the corresponding idealized B-matrix, passed as input to the minimization procedure.

Figure 7.49 on page 277 provides an example of a trial for S^1 in E^2 with $n = 200$ vertices and shows the progressive relocation of monads from their initial randomized locations over the course of ~ 1000 iterations. Figure 7.50 shows the corresponding analysis. In 7.50(a) the initial coordinates were obtained by assigning random x and y values from a uniform distribution in the range $(-1.2, 1.2)$ – thus the linear increase in radial density up to the maximum radius and subsequent drop off. Figure 7.50(b) shows the cumulative effect of micro-adjustments to coordinates according to the sum of the spring forces experienced by each monad, effectively lowering the 'energy' of the system, which falls rapidly over the first ~ 300 iterations as the vertices are attracted toward the centre-of-mass position (c/f figure 7.49(c)); the steep descent is followed by a gradual further decline in energy as the S^1 shape 'inflates', leading finally to the

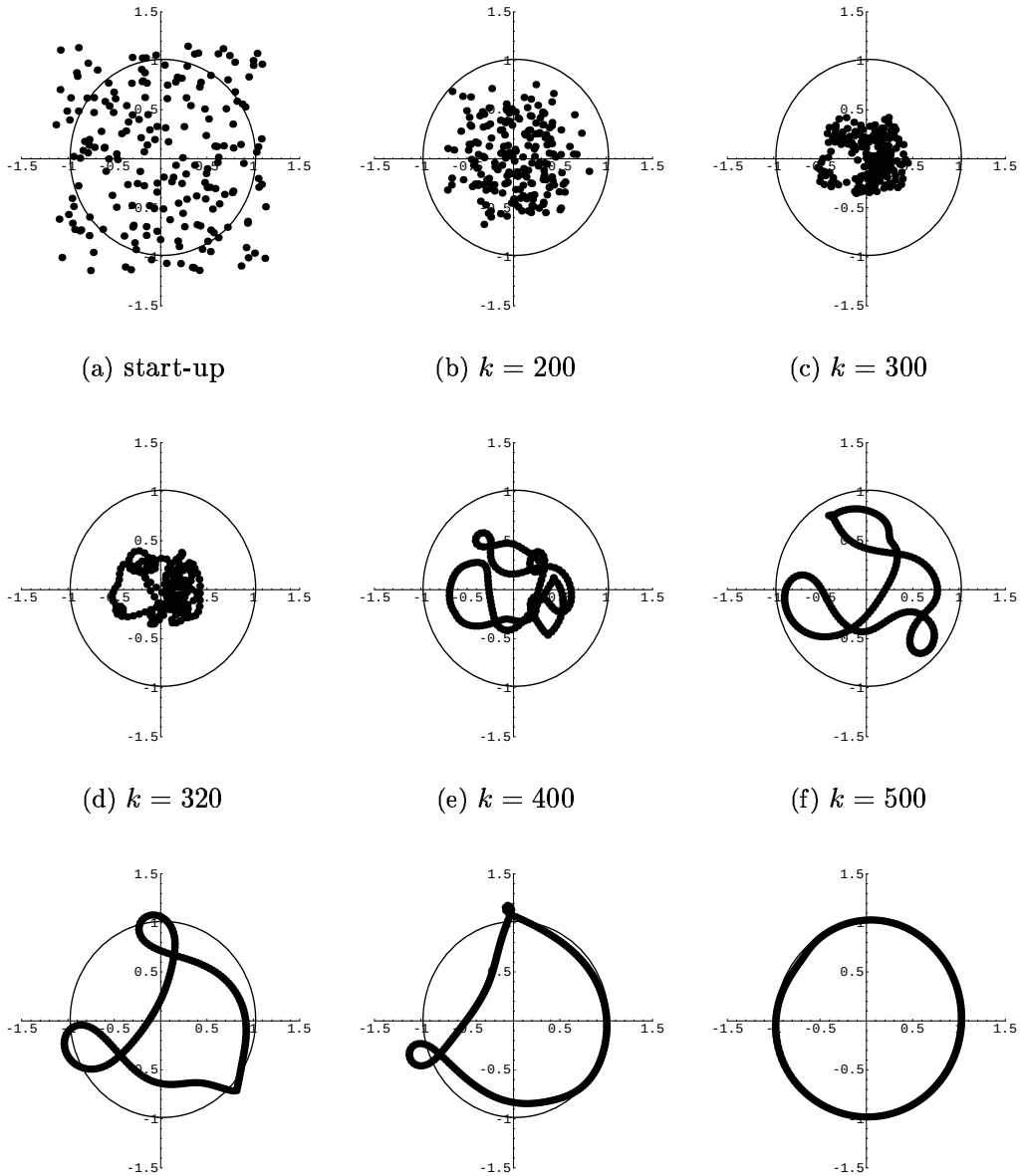
signature histogram of figure 7.50(c), where all nodes share a common radius from the centre-of-mass position.

Figures 7.51 and 7.52 on pp.279, 280 show the corresponding results from a trial for S^2 in E^3 with $n = 418$ monads. Because of the greater number of monads required for a reasonable density of point in E^3, and hence the greater computational overhead, the repositioning of monads was effected in larger increments than the previous trial and the energy minimization thus achieved within 125 iterations. In this trial, the random assignment of coordinates was undertaken in spherical form. Radii were drawn from a uniform distribution in the range $(0, 2)$ and angles θ and ϕ drawn from uniform distributions in the ranges $(0, 2\pi)$ and $(0, \pi)$ respectively, resulting in the relatively flat distribution of radii shown in figure 7.52(a). As the figures show, the coarser adjustments were not detrimental to the success of the procedure and, again, a clear signature for S^2 in E^3 was obtained (see figure 7.52(c)).

Figures 7.53 and 7.54 on pp.281, 282 provide results from a trial for S^3 in E^4 with $n = 800$ monads. As with the lower-dimensional cases, the single radial value in the final histogram reveals the desired signature, corresponding to the presence of S^3 in this case, of course.

Minimization procedure results on actual gebits

Figures 7.55 and 7.56 (on pages 283 and 284) show typical results when the minimization procedure is applied to HPS B-matrices generated by the iterator and extracted after running the system beyond the start-up phase. Figure 7.55 obtained from a modest set of 100 monads after running the iterative map with the SRN term biased to produce a greater number of large B_{ij}. Minimization was then undertaken with

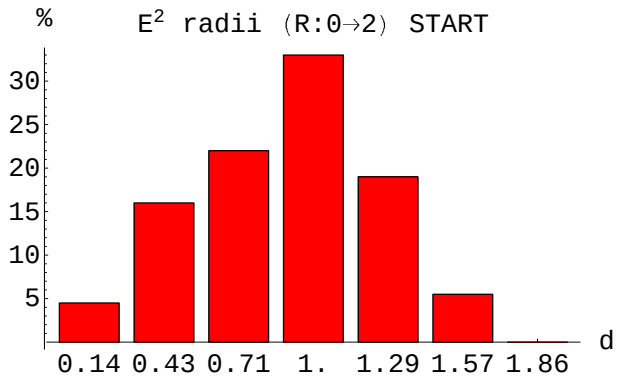


(a) start-up (b) $k = 200$ (c) $k = 300$

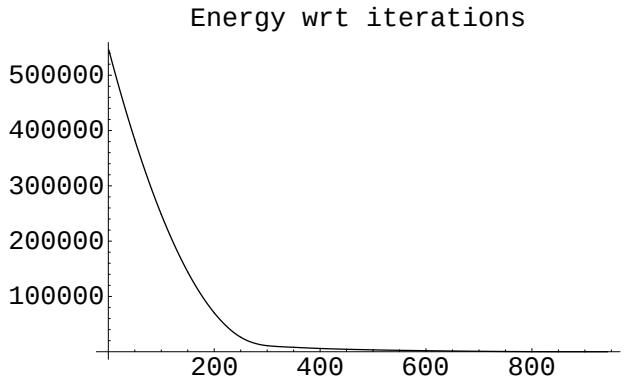
(d) $k = 320$ (e) $k = 400$ (f) $k = 500$

(g) $k = 600$ (h) $k = 700$ (i) $k = 940$

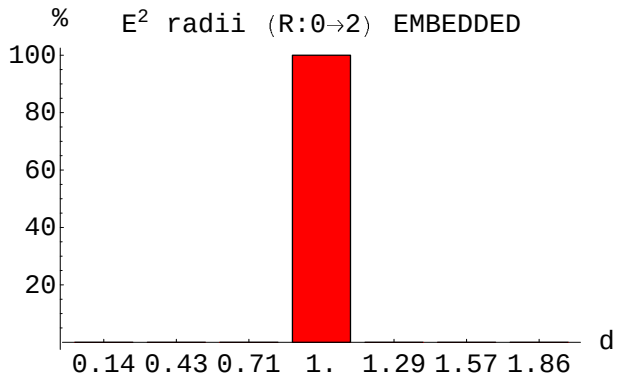
Figure 7.49: Trial embedding: S^1 in E^2
showing how the randomly placed nodes are progressively repositioned by the energy minimization procedure according to the information content of the B-matrix, which was constructed explicitly to contain an unambiguous S^1 signature – the 'target' shape represented by the lighter circle.



(a)

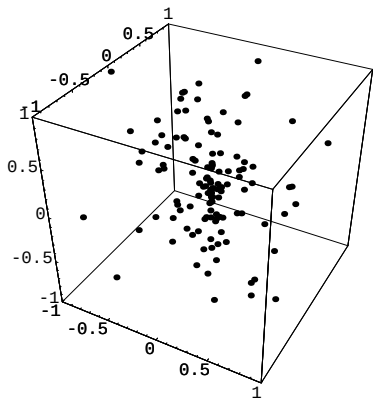


(b)

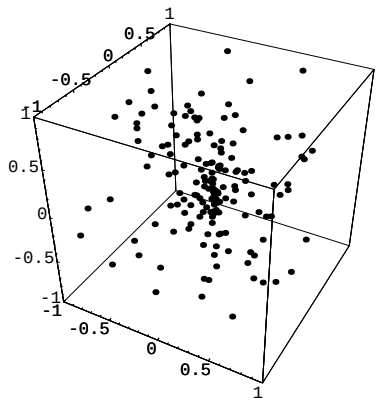


(c)

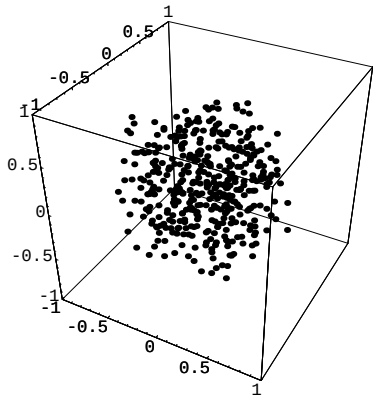
Figure 7.50: Trial embedding: S^1 in E^2 – analysis
(a) is a relative frequency histogram of radii from the initial randomized coordinates; (b) shows the effect of the energy minimization with respect to each iteration of the procedure; (c) is a relative frequency histogram of the final radii – the single radial value demonstrates the 'ideal' signature of S^1 in E^2.



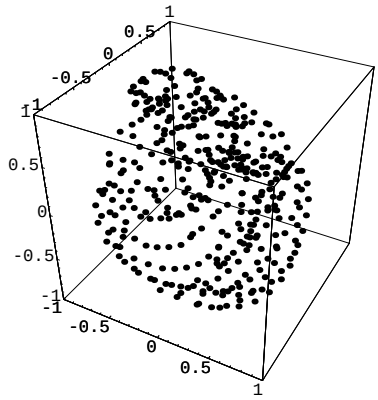
(a) start-up



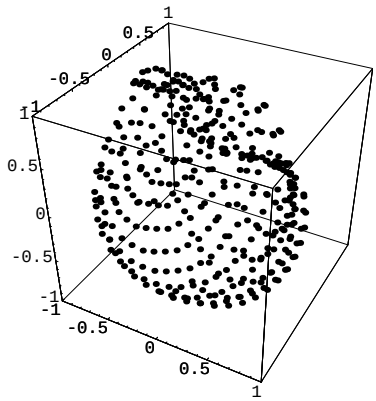
(b) $k = 20$



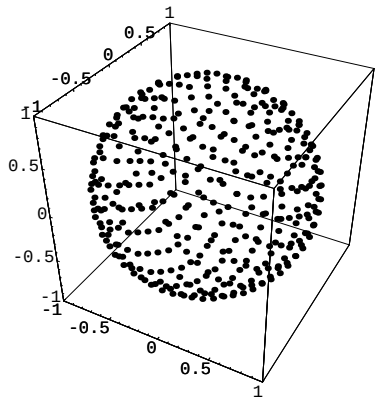
(c) $k = 70$



(d) $k = 100$

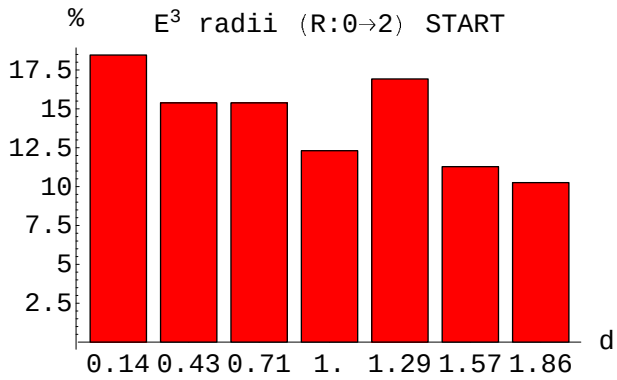


(e) $k = 110$

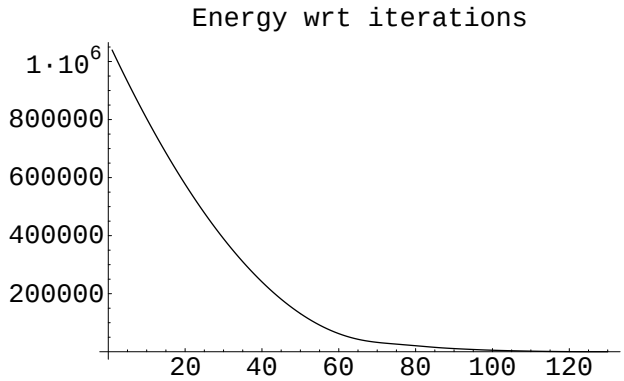


(f) $k = 125$

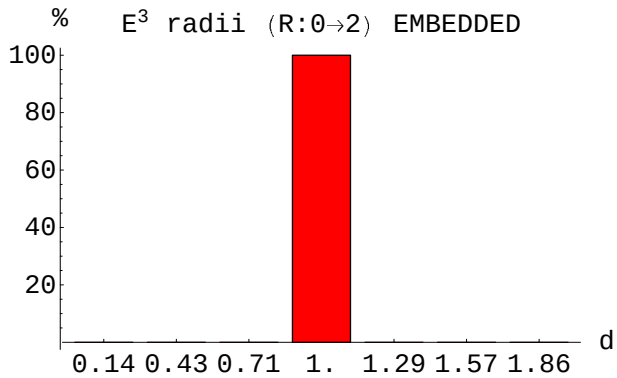
Figure 7.51: Trial embedding: S^2 in E^3



(a)

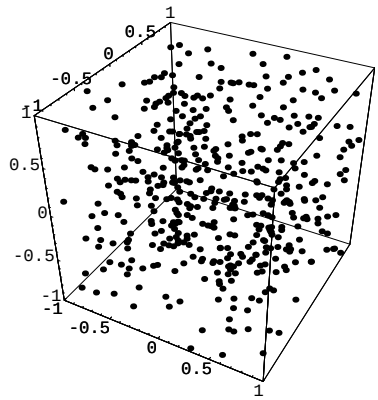


(b)

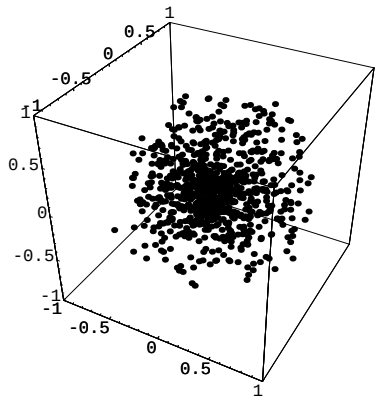


(c)

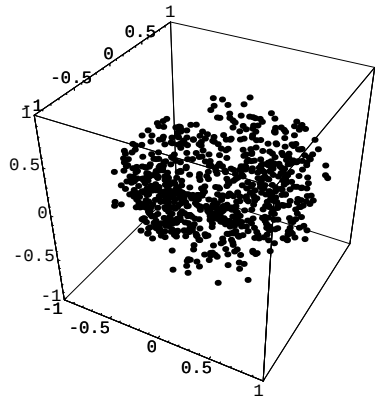
Figure 7.52: Trial embedding: S^2 in E^3 – analysis
(a) is a relative frequency histogram of radii from the initial randomized coordinates; (b) shows the effect of the energy minimization with respect to each iteration of the procedure; (c) is a relative frequency histogram of the final radii – the single radial value demonstrates the 'ideal' signature of S^2 in E^3.



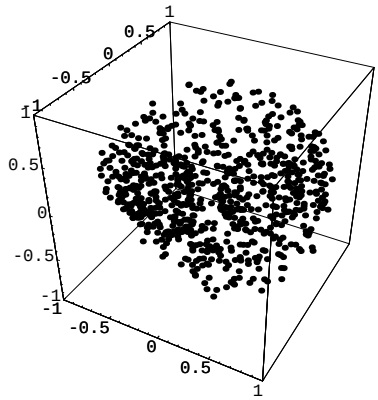
(a) start-up



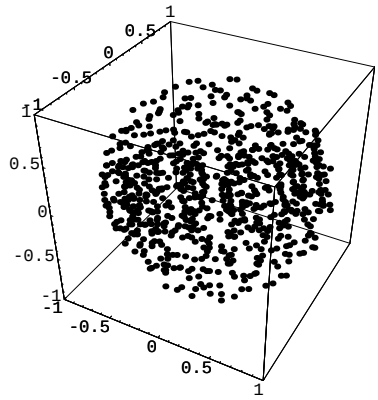
(b) $k = 35$



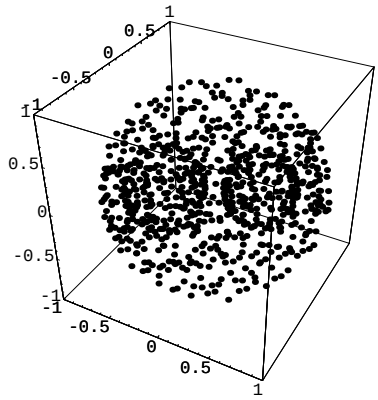
(c) $k = 60$



(d) $k = 90$

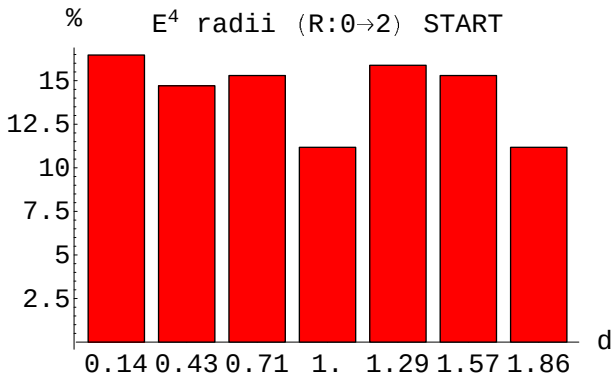


(e) $k = 110$

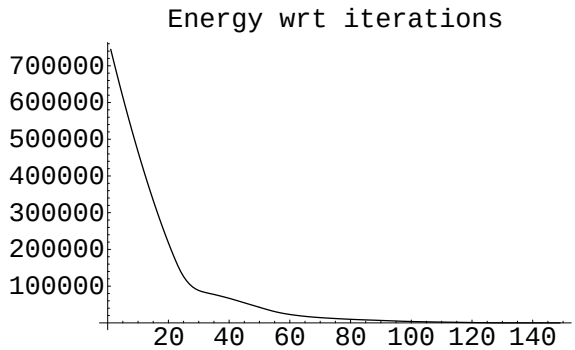


(f) $k = 145$

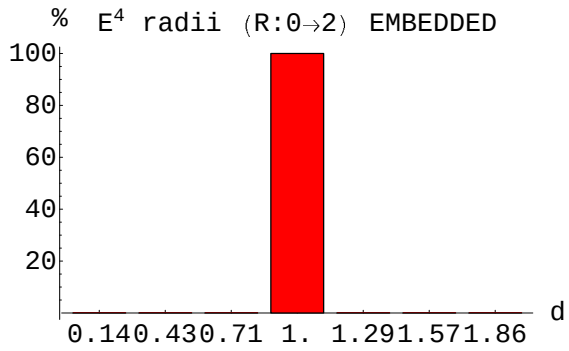
Figure 7.53: Trial embedding: S^3 in E^4
with fourth coordinate suppressed to permit plotting. The hyperspherical geometry is indicated by the 'sphere within a sphere' roughly visible in the final projection, (f).



(a)



(b)



(c)

Figure 7.54: Trial embedding: S^3 in E^4 – analysis
(a) is a relative frequency histogram of radii from the initial randomized coordinates; (b) shows the effect of the energy minimization with respect to each iteration of the procedure; (c) is a relative frequency histogram of the final radii – the single radial value demonstrates the 'ideal' signature of S^3 in E^4.

respect to E^2, E^3, and E^4.

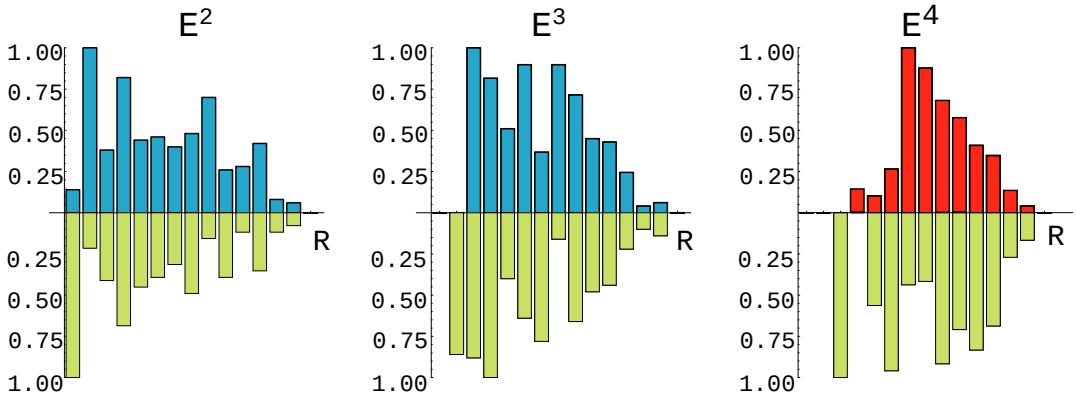
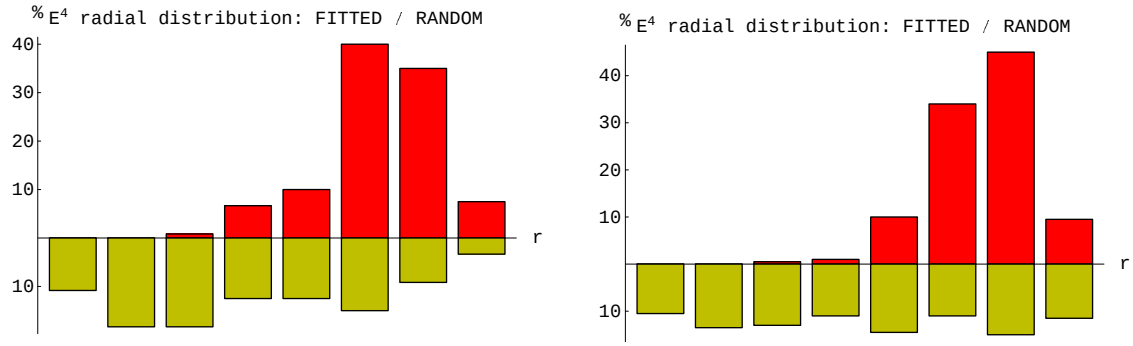


Figure 7.55: Gebit embeddings: S^3 in E^n, $n = 2, 3, 4$
showing density of monads plotted above the horizontal axis against radial distance; plots below the axis show the randomly generated cases for the same gebit. Note the clear peak and diminishing right tail for the E^4 case demonstrating the presence in B of S^3 structure with strong yet fuzzy embeddability whereas the E^2 and E^3 cases have relatively 'flat' radial distributions, essentially no different from the random embeddings.

The key result from this study is the signature peak for the radial distribution in the E^4 case, showing the presence of an intrinsic S^3 structure, in contrast to the distributions in the E^2 and E^3 cases, which are essentially no different from the distributions of the randomized embeddings. The finite width of the peak in the E^4 distribution indicates links that are outside the S^3 geometry and thus non-local with respect to that geometry. The extent of that departure is, however, extreme in this instance due to the bias introduced to the SRN.

Figures 7.56, 7.56, and 7.58 give results from three further trials that looked solely at the E^4 case with $n = 120$, 200, and 1000 respectively. In these, the B-matrices obtained after 1000 iterations of the HPS systems and with no bias given to the SRN

term. Generally, the signature is stronger and more consistent for larger gebits, yet still indicating 'fuzzy embeddability' by the non-zero density for neighbouring radii. In figure 7.56, the histograms again show the density of monads plotted above the horizontal axis against radial distance, with randomly generated cases for the same gebits plotted below the axis for comparison.



(a) $n = 120$ (b) $n = 200$

Figure 7.56: Gebit embeddings: S^3 in E^4
again showing density of monads plotted above the horizontal axis against radial distance, with randomly generated cases for the same gebits plotted below the axis for comparison.

The randomized case is omitted in figure 7.57 in favour of showing the radial scale. In particular, note the very strong S^3 signature indicated by the dominant peak at $r = L/2 \sim 12$, consistent with the theoretical $L = 26$ for $n = 1000$. Figure 7.58 shows the energy minimization for the last example. Equilibrium was reached when the nett forces had gone to ~ 0 yet the final energy was non-zero, indicating that the S^3 geometry remains under some residual tension.

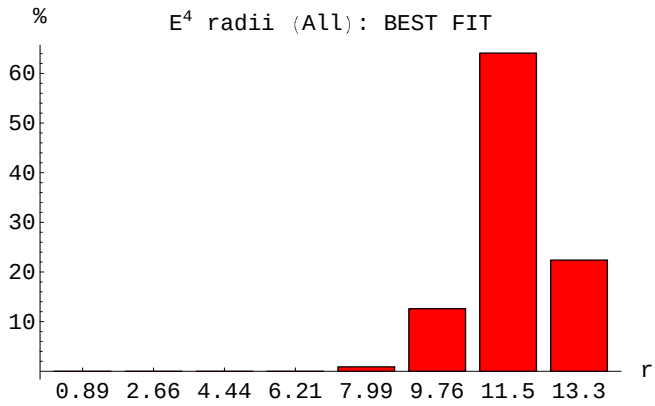


Figure 7.57: Gebit embeddings: S^3 in E^4, $n = 1000$
In particular, note the very strong S^3 signature indicated by the dominant peak at $r = L/2 \sim 12$, consistent with the theoretical $L = 26$ for $n = 1000$.

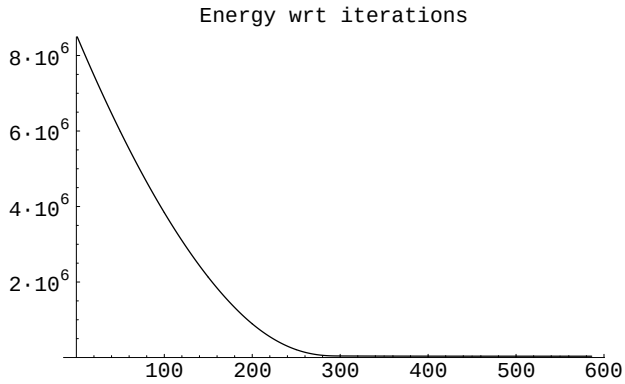


Figure 7.58: Energy minimization: S^3 in E^4
pertaining to the previous figure 7.57 with 1000 monads. The initial 'energy' was $\sim 8.5 \times 10^6$ and the final energy ~ 35000 with the system in equilibrium after nett forces have gone to ~ 0.

7.4 Concluding remarks for HPS

HPS theory began where HQS theory left off. Utilizing recursion and stochasticity via a form of Wiener process, the HPS *iterator* of equation 7.2.1 generates successive HQS-like states, constructed from abstract entities dubbed *monads* via pair-wise relations driven by the *SRN* and the behaviour of the *binder* term in the iterator. These successive states rise above the fundamental condensate to an equilibrium plateau where order and disorder combine to synthesize dynamical structure in the form of *gebits* – networks of information possessing an intrinsic and inevitable dominant geometrical property: the signature of a closed, three-dimensional space, the hypersphere S^3. The spontaneous appearance of such structure is entirely relational; it does not arise within any *a priori* geometrical background yet it introduces an endophysical notion of locality in the sense of having *position*. By construction, it is the most robust structure but with the further, richer quality that within the gebits lesser and more transient relations introduce non-local connections that distort and diffuse the S^3 embedding, making it 'fuzzy' and imbued with softer emergent quantum-like modes of behaviour. In the behaviour of this iterative process, there are striking parallels with 'natural hermeneutics', as defined by Kampis [362]:

> The most fundamental ingredients of a hermeneutic scheme will be postulated as this (and nothing else):
>
> - there is an iterative unfolding of information content and a subsequent continual change of what is already unfolded;
> - there is an historical element in this process;

- the key elements of the situation have a qualitative rather than quantitative nature;
- a certain degree of circularity is present.

It should be clear that a HPS exhibits each of these characteristics. The parallels will be seen to extend even further as HPS theory is developed into the theory of *Process Physics* (see Chapter 8).

7.4.1 Topological defects and fuzzy homotopy

The mechanism for this has not been discussed explicitly up to this point, but it follows both intuitively and inductively from the various analyses and results that have been described thus far: by coarse-graining the B-matrix at an appropriate threshold, one is left with a matrix B' the corresponding graph of which, when its depth structure is analyzed, is such that the most likely arrangement of vertices in a connected component follows a $\sin^2$ distribution (per §6.6.2). In earlier discussion pertaining to catastrophic events and the threshold problem, it was indicated (§7.3.1 on page 221) that a sub-maximal cut would resolve the threshold problem while suggesting fractal structure, implying that an extreme value is merely an instance of a higher level set the other members of which reside outside the sample under consideration. Thus the large components of the sub-maximal cut may be seen as just one of many classes having the property of sparse large (relative to that class) links. Inductively, then, a still lesser cut may be taken and so provide a lower level class that likewise possesses the same universal property of sparse (relatively) large entries and thus satisfies the conditions required for the $\sin^2$ result. Hence the lesser and more transient links within a gebit have, themselves, the connectivity and embedding

geometry of all gebits (since this is an intrinsic property). But, in the context of what might be called the 'base' gebit, the adjacency relations so defined will, with near certainty, be inconsistent with those of the 'parent' – that is, they have a fuzzy and non-local embedding in the base gebit. To put it another way, these extra links act as topological defects that provide an indication of fuzzy homotopies[4] The results on embeddability testing presented in the previous section lend support to this picture by the appearance of a dominant peak in the radial distribution of monads occasioned by large-valued link variables and lesser adjacent peaks occasioned by the presence of weaker links.

7.4.2 Interpreting monads, 'bootstrapping', and the end-game

The base gebits $g_1, g_2, \ldots$ arising from the SRN together with their embedded topological defects have another remarkable property: the binder mechanism responds vigorously to counter singular or degenerate tendencies in the network so that preferential treatment is given to monads associated with the (potential) kernel – what have been referred to as *active nodes* – and new large values are generated by the binder to cross-link new and existing gebits (see §7.3.2, from page 257). It was shown in §7.3.3 on page 268 that the distribution of pair-wise distances between active nodes is not significantly different from the distribution between all nodes and thus the new links created by the binder are distributed randomly across the network. Since this cross-linking is random, base gebits might be treated as abstract unit entities themselves, that is, without regard for their internal monadic constituents, and the whole analysis repeated from this new perspective. One must then conclude that the resulting

[4]Simply put, a homotopy is an embedding of one space into another.

network of base gebits has, itself, the same intrinsic geometrical property and forms a higher-level S^3 with embedded topological defects. Now, finally, it becomes possible to do what could not be done in §6.6 – to give meaning to, and better define, monads via their functional relationship to this larger scheme rather than by objective definition: for, if gebits may be viewed and treated as if they were monads, then monads may be viewed and treated as if they were, themselves, networks of connected gebits. That is, the distinction is entirely arbitrary – the functional relationships between units are identical and it becomes a matter of choice to employ the label 'gebit' when one wishes to acknowledge interior relations, or the label 'monad' otherwise. For this reason, an HPS, being a process that is both self-initiating and self-sustaining, is a *bootstrap* system (i.e. the start-up components and emergent components are identical in form). Here, then, is further support for the fractal nature of HPS, a required condition for hiding, via universality, the details of the particular HPS realization. Moreover, it may now be seen that the HPS scheme resolves the 'end-game' problem because there are neither lower nor upper levels to impose bounds: the act of investigating the system is akin to cutting a circle – the location of the cut is an arbitrary and externally imposed condition that, while serving the investigator, is yet independent of the system under examination.

7.4.3 Process space and the essence of time

In the course of ongoing iterations, the existing gebits acquire more cross-links until their latent null space has diminished to the point where they escape the attention of the binder mechanism, losing their ability to undergo further linking. That is, they eventually lose their 'stickiness' and decay. Hence, the space defined by HPS

theory is 3-D, with the characteristics of S^3, but the structure is independent of its component gebits, which are continually undergoing replacement: in stark contrast with static continuum geometrical spaces, the emergent space of HPS theory is dynamic and fractal – HPS realizes what Kampis terms a 'self-modifying network' [363], whereby the self-modification of the system occurs as a consequence of the higher organization (here, *self*-organization) that integrates its constituent parts. This notion of an emergent fractal process space is depicted by the artist's impression presented in figure 7.59[5]. It is a dissipative relational information *process* space driven by the noise of self-referencing. If the SRN is turned off, the emergent structure subsides into the condensate and ceases to exist.

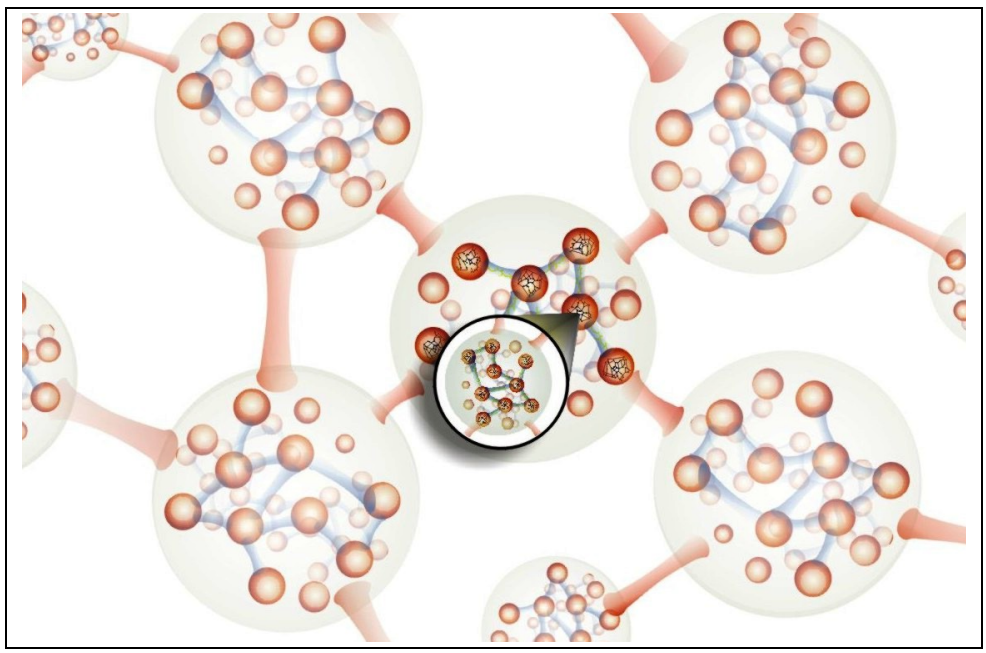


Figure 7.59: Artist's impression: fractal process space
with linked and embedded gebits forming a 3-D fractal process space characterized by a 'quantum foam' behaviour. The magnifying glass view, in which monads are seen as gebits and *vice versa*, emphasizes the fractal nature of the network.

[5]This picture, by the author, was first published as the cover graphic of [364].

The fundamental and essential insight that brought about the development of HPS theory from the precursor HQS theory was the recognition that time is fundamental, that there is nothing *like* time with which to construct a model such that the phenomenology of time is an emergent property. The introduction of process and the stochasticity of self-referential noise imbues HPS theory with *vitae essentia* both by the creative spontaneous generation of spatial structure and by the capture of what may be called the 'present moment effect'. In §7.3.2, figures 7.30 to 7.35 demonstrated how the history of the system could be recorded and replayed precisely, but if one intervenes at some point in that history then, even in the presence of identical SRN values, the effects of that intervention ripple forward as the binder responds as an active and creative generator of reactive phenomena that propagate along the forward time line. While it is important to note that the iterations of the map do not constitute, *a priori*, the phenomenon of time (their function is to produce the needed fractal structure), HPS theory nevertheless captures the essence of empirical or experiential time in two distinct ways: firstly as a consequence of its stochastic properties – there is an inescapable 'arrow of time' that points inexorably from that which has gone before to that which is yet to be and one cannot run the system in reverse to recover an earlier state because the map is unidirectional (owing to the noise term, there is no way to define an inverse mapping); and secondly as a consequence of both the SRN and the non-local action of the binder – the system evolves a history (a recordable past) that, in principle, may be broadly inferred from the presence of persistent relational forms, in which sense the mapping shows a natural partial memory phenomenon, but the present moment is entirely contingent on both the SRN and the specific detail of that history, while the future awaits creation – it

is completely unknowable without performing the iterations.[6]

HPS theory demonstrates how the nature of space and the nature of time are deeply related to the 'logic of the limitations of logic' [364] and that, contrary to convention, both space and time are to be understood as *process* rather than labelled and treated as objects. Yet endophysical experiential time is only predicted in HPS theory if there is an emergent ordered sequencing of events at the level of universality – that is, above the level of details that are purely incidental to a particular realization: it might be that the iterative map fails to produce change at a high level, in which case there would be no intrinsic temporal phenomenology, a view consistent with, for example, that expressed by Kampis: "the new ontology suggested by the 'endo' logic ... acknowledges that every moment of time has a real value in that it produces, or may produce, something genuinely new, using the already available elements in a novel relation" [171]. To proceed, and to demonstrate the viability of the HPS model of reality, it is necessary to consider how such a system generates emergent quantum-like behaviour, including the important but often ignored process of *objectification*.

[6]This is very unlike the conventional time-symmetric differential equations of traditional physics, which is not to say that the convention in physics of modelling time as a geometric construct is not a useful device but, rather, to suggest that the simplicity and appeal of the geometric model of time may detract from a full appreciation of the richness of that which is being modelled. Moreover, the geometric model seductively admits a regress of temporal arguments and paradoxes, spawning much speculation and debate that does not appear to have been fruitful in furthering our understanding of the universe.

Chapter 8

Process Physics

We are The Matrix. We create Order from Chaos.

– The Matrix, Gargoyles: Episode 46

8.1 Introduction

Figuratively, the HQS formulation provided the building blocks of a nascent universe, while HPS theory provided a mechanism by which those building blocks – emergent relational structures dubbed *gebits* – arise and self-assemble to form a three-dimensional fractal process space. Both analytical and numerical studies have indicated that this process space, driven by SRN, is a complex dynamical network that expands rapidly and self-organizes, behaving essentially as a dynamical Prigogine-like dissipative order-disorder system [66] that is now to be associated with the phenomenon of cosmological space.

This chapter outlines and sketches the further developments that extend HPS theory to what is now known as *Process Physics* [364], the fundamental nature and

richness of which is such that a fully detailed and complete account lies well beyond the scope of the present work. The aim here is to indicate, as a first step, the key concepts whereby the theory connects with, and motivates, the central higher level descriptions of the standard model of particle physics and of gravitation, and to also illustrate the potential for *Process Physics* to reach beyond those models.

8.2 Modelling gebits and their topological defects

In [18], Cahill described the development of a Quantum Homotopic Field Theory (QHFT) for *Process Physics* by first considering the special case of the iterator when the SRN is artificially fixed at a particular w. Then, the iterator is equivalent to the minimization of 'energy' expression,

$$E[B; w] = \frac{\alpha}{2}\mathrm{Tr}[B^2] - \alpha\mathrm{TrLn}[B] + \mathrm{Tr}[wB]. \tag{8.2.1}$$

For disconnected gebits g_i, g_j this energy is additve, so that $E[g_i \oplus g_j] = E[g_i] + E[g_j]$. Suppose now that the fixed w has the form of a single disconnected gebit approximating an S^3 network but containing one embedded defect that also has the S^3 signature: that is, the frozen SRN gebit is decomposed into base gebit, defect gebit, and the linkings between the two. The rest of the network may be ignored since this gebit is disconnected. In the case that $\det(w) \gg 0$ then this gebit is non-sticky and, for small α, the iterator will converge to $B \approx \frac{1}{\alpha}w$, merely an enhancement of the gebit. On the other hand, if $\det(w)$ is small – the most likely scenario, since the most probable configuration is that of rare large w_{ij} forming tree structures within a background of lesser valued w_{ij} – then the gebit attracts the attention of the binder via the inverse, B^{-1}, and new links are generated to counter degenerate tendencies,

with the consequence that the embedded defect 'relaxes' with respect to the base gebit and the energy measure. Such relaxation is an example of a non-linear elastic process [365]. The gebit has the form of a mapping from a base space to a target space, $\pi : S \to \Sigma$. The continuum form for the elastic energy of such an embedding was constructed by Manton [366, 367] and for $\pi : S^3 \to S^3$ it is the Skyrme energy

$$E[U] = \int \left[-\frac{1}{2}\mathrm{Tr}\left(\partial_i U U^{-1} \partial_i U U^{-1}\right) - \frac{1}{16}\mathrm{Tr}\left[\partial_i U U^{-1}, \partial_i U U^{-1}\right]^2 \right], \qquad (8.2.2)$$

where $U \equiv U(x)$ is an element of $SU(2)$. Via $U(x) = \sigma(x) + i\vec{\pi}(x)\cdot\vec{\tau}$, where the τ_i are Pauli matrices, $\sigma(x)^2 + \vec{\pi}(x)^2 = 1$ parameterizes S^3 as a unit hypersphere embedded in E^4. Non-trivial minima of $E[U]$ are a particular form of topological solitons, called Skyrmions, having $Z = \pm 1, \pm 2, \ldots$, where Z is the winding number of the mapping, given by

$$Z = \frac{1}{24\pi^2} \int \sum \varepsilon_{ijk} \mathrm{Tr}\left(\partial_i U U^{-1} \partial_j U U^{-1} \partial_k U U^{-1}\right). \qquad (8.2.3)$$

Cahill points out that "the first key to extracting emergent phenomena ... is the validity of this continuum analogue, namely that $E[B; w]$ and $E[U]$ are describing essentially the same 'energy' reduction process. This should be amenable to detailed analysis." [18]

The time evolution of the full iterator is, of course, far richer than this analysis with fixed SRN suggests. When driven by ongoing noise, the iterator produces an endless flux of pattern creation, linkage, and then, for most patterns, subsequent decay[1]. The task is to identify particular patterns that are resilient, surviving the flux despite the transience of their constituent components, and then to attempt to categorize their modes of behaviour. The 'giant components' described in §7.3.1 on page 223 are examples of such resilient, or persistent, patterns and there is some

[1] "Ta panta re – all is flux" (Heraclitus).

indication that they may drift through the system, even displaying ergodic tendencies with respect to the available monadic resources. From HPS theory, gebits emerging from the SRN have active nodes and embedded topological defects, which themselves have active nodes. From the fractal nature of the system, even the defects will likely possess embedded defects and *all* of the defects, whatever their classification, will have associated winding numbers. The energy analogy suggests that defects at the same fractal depth and possessing opposite winding numbers may annihilate by drifting together and merging. Furthermore, consistent with the fuzzy embeddability that appears to be a generic property of the system, it seems unlikely that the embedding of defects would be 'classical' in the sense of being described by a mapping $\pi(x)$ but, rather, would also exhibit fuzziness requiring description by some functional form $F[\pi]$ that would correspond to a classical embedding only if $F[\pi]$ possesses a supremum at a specific $\pi \equiv \pi_{\text{classical}}$.

The emergent hierarchical fractal behaviour of HPS theory is a generic property that is suggestive of a quantum space foam, but containing topological defects that will be preserved unless annihilation events occur. If the defects are sufficiently rich in fractal structure as to be preserved (one might think of these as 'knots') then their formation most likely occurred as the process space arose from its initial, essentially random, state[2]. Once trapped, such defects (knots) must inevitably undergo constant renewal by self-replication – the process of replacing constituent components while retaining the characteristic pattern[3] – and so traverse the greater process space. The

[2]One might speculate that this phase corresponds, at least analogously, to early stages of the Big Bang.

[3]Metphorically, consider a closed loop of knotted string: as the string ages, replace small sections with new string; eventually, all of the string will be replaced yet the structure of the loop and embedded knot remain – the topology of the system, the pattern, is preserved though its constituents be transient. There are many such metaphors in complex systems: eddies in a river; patterns of

conjecture is that these residual preserved topological defects correspond to matter in a very fundamental sense, most likely being sub-quantum in the form of preons. Moreover, the same schema determines the behaviour both of the process space and its defects (in a real sense, the distinction is artificial, albeit conceptually convenient) and so the further conjecture is that *Process Physics* provides a unification of space and matter phenomenology in a manner that suggests an appropriate generic descriptor for higher-level network activity is the QHFT introduced in the next section.

8.3 Quantum Homotopic Field Theory

SRN is the dominant resource available to the iterator. Large-valued w_{ij} provide 'sticky' gebits to be assembled, via the binder mechanism, into the 3-D process space. Topological defects within these gebits, and within the process space as a whole, require a subtle description. Recall that the most significant behaviour for those defects that are sufficiently large (in terms of the number of constituent components), and topologically non-trivial, is that of *persistence*, which manifests by topological self-replication. As indicated in the previous section, continuum homotopy theory may be applied to identify non-trivial topology. QHFT is a form of quantum field theory developed to track the coarse-grained behaviour of such a system by modelling both the process space and the defects. It has the form of a functional Schrödinger equation for the discrete time evolution of a wave-functional Ψ:

$$\Psi[..., \pi_{\alpha\beta}, ..., t + \Delta t] = \Psi[..., \pi_{\alpha\beta}, ..., t] - iH\Psi[..., \pi_{\alpha\beta}, ..., t]\Delta t + \text{ QSD}^4 terms, \quad (8.3.1)$$

traffic flow; flocking and herding behaviour; social organizations; even the human body – the essence of a given individual persists though the cells of the body undergo continual replacement.

where the configuration space is provided by the set of all possible homotopic mappings so that $\pi_{\alpha\beta}$ maps from S_β to S_α, with $S_\gamma \in \{S_1, S_2, S_3, \ldots\}$, the set of all possible gebits (the topological defects are not necessarily of the S^3 form). Since this configuration space is prescribed by the fractal depth of gebits under consideration, the time step Δt in (8.3.1) is to be chosen relative to the scale of the fractal processes being explicitly described (i.e. at smaller scales, smaller Δt would be required), invoking a (finite) renormalization scheme. In conventional terminology and in the absence of the QSD term, equation 8.3.1 represents a 'third quantized' system [368], considered as perhaps capable of generating a quantum theory of gravity. In Process Physics, this is the emergent behaviour of the network and the conjecture is that it does, indeed, lead to quantum gravity and also quantum matter. Moreover, the origin of (8.3.1) is understood and it will lead first to quantum gravity and then to the classical theory as a higher level abstraction, rather than attempting the deeper theory via some heuristic quantization procedure.

Depending on the maxima of Ψ and the connectivity of the resulting dominant mappings, such mappings may be interpreted either as embeddings or as links; in this sense, revisiting figure 7.59 on page 290 reveals the picture to be indicative of the dominant process space form within Ψ and showing both embeddings and links; the original configuration of nodes and links at start-up is irrecoverable. Of course, the picture is purely for illustrative purposes and must not be taken too literally – the process space cannot actually be embedded in a finite dimensional geometric space while preserving the emergent metric since it is composed of infinitely nested (fractal) finite-dimensional closed spaces.

[4] 'QSD': Quantum State Diffusion; see 8.3.1 on page 300.

In the previous section, it was proposed that Manton's non-linear elasticity interpretation of the Skyrme energy provides an appropriate form for the analysis. This then suggests that the hamiltonian H should be the functional operator

$$H = \sum_{\alpha \neq \beta} h \left[\frac{\partial}{\partial \pi_{\alpha\beta}}, \pi_{\alpha\beta} \right], \tag{8.3.2}$$

where $h[\frac{\partial}{\partial \pi}, \pi]$ is the quantum Skyrme hamiltonian operator for the system, based on making fuzzy the mappings $\pi : S \to \Sigma$ by having h act on wave-functionals of the form $\Psi[\pi(x); t]$. Then H is the sum of pairwise embeddings or homotopy hamiltonians and the corresponding functional Schrödinger equation would describe the time evolution of quantized Skyrmions in a fixed base space and with $\Sigma \in SU(2)$.

Again from [18], Cahill showed that, in the absence of the QSD terms, the time evolution in equation 8.3.1 can be written formally as a functional integral

$$\Psi\left[\{\pi_{\alpha\beta}\}; t'\right] = \int \prod_{\alpha \neq \beta} \mathcal{D}\tilde{\pi}_{\alpha\beta}\, e^{iS\left[\{\tilde{\pi}_{\alpha\beta}\}\right]}\, \Psi\left[\{\pi_{\alpha\beta}\}; t\right], \tag{8.3.3}$$

where the action is the sum of pairwise actions

$$S\left[\{\tilde{\pi}_{\alpha\beta}\}\right] = \sum_{\alpha \neq \beta} S_{\alpha\beta}\left[\tilde{\pi}_{\alpha\beta}\right], \tag{8.3.4}$$

$$\begin{aligned} S_{\alpha\beta}\left[\tilde{\pi}\right] = \int_t^{t'} dt'' \int d^n x \sqrt{-g} \Bigg[& \frac{1}{2} \mathrm{Tr}(\partial_\mu \tilde{U}\tilde{U}^{-1} \partial^\mu \tilde{U}\tilde{U}^{-1}) \\ & + \frac{1}{16} \mathrm{Tr}\left[\partial_\mu \tilde{U}\tilde{U}^{-1},\, \partial^\nu \tilde{U}\tilde{U}^{-1}\right]^2 \Bigg], \end{aligned} \tag{8.3.5}$$

and the (now) time-dependent (indicated by the tilde symbol, per the continuum limit notation) mappings $\tilde{\pi}$ are parameterized by $\tilde{U}(x,t)$, with $\tilde{U} \in S_\alpha$. Each homotopic mapping appears in both orders, namely $\pi_{\alpha\beta}$ and $\pi_{\beta\alpha}$, and the metric $g_{\mu\nu}$ is that of the n-dimensional base space, S_β, in $\pi_{\alpha\beta} : S_\beta \to S_\alpha$. The functional derivatives in

the quantum hamiltonian of equation 8.3.2 now manifest as the time components ∂_0 in equation 8.3.5, as is usual in the functional integral formalism, and so now (8.3.5) has the form of a 'classical' action, providing the emergence of classical fields, though the emergence of classical phenomena is a more complicated process. Equations 8.3.1 and 8.3.5 describe an infinite set of quantum Skyrme systems coupled in a pairwise manner.

8.3.1 Quantum State Diffusion

Quantum State Diffusion (QSD) [127, 129] theory is an alternative quantum theory from which classical dynamics, quantum dynamics and localization dynamics are derived. Dissipation of coherence is distinguished from destruction of coherence – the state diffusion destroys coherence and produces the localization or reduction that makes classical dynamics possible. The theory is a development of the environmental quantum state diffusion theory of Gisin and Percival [76] and particularly resembles earlier proposals by Gisin and by Milburn. It is also related to the spontaneous localization theories of Ghirardi, Rimini and Weber [67], of Diosi [68] and of Pearle [125, 126]. Diosi, Gisin, Halliwell, and Percival [131] demonstrated a close connection between the decoherent histories approach to quantum mechanics and the quantum state diffusion picture, for open quantum systems described by a master equation of Lindblad form.

The additional QSD terms in equation 8.3.1 involve summation over the class of

Linblad[5] functional operators L_γ, specifically [18]:

$$\text{QSD terms} = \sum_\gamma \left(<L_\gamma^\dagger> L_\gamma - \frac{1}{2} L_\gamma^\dagger L_\gamma - <L_\gamma^\dagger><L_\gamma> \right) \Psi \Delta t + \sum_\gamma \left(L_\gamma - <L_\gamma> \right) \Psi \Delta \xi_\gamma, \tag{8.3.6}$$

where the $\Delta\xi_\gamma$ are complex statistical variables with means $M(\Delta\xi_\gamma)$, $M(\Delta\xi_\gamma \Delta\xi_{\gamma'}) = 0$ and $M(\Delta\xi_\gamma^{(}*)\Delta\xi_{\gamma'}) = \partial(\gamma - \gamma')\Delta t$. The QSD terms are up to 5th-order in Ψ as, in general,

$$<A>_t = \int \prod_{\alpha \neq \beta} \mathcal{D}\pi_{\alpha\beta} \Psi \left[\{\pi_{\alpha\beta}\}; t \right]^* A \Psi \left[\{\pi_{\alpha\beta}\}; t \right], \tag{8.3.7}$$

and have the remarkable property that the unitary time evolution of equation 8.3.1 is maintained. These QSD terms are non-linear, stochastic, and ultimately responsible for the emergence of classicality via an *objectification* process but, in particular, they result in the collapse of the wave function(al) during quantum measurements by tending to 'focus' the fuzzy homotopies towards 'sharp' classical homotopies.

The formalism and interpretation of decoherence usually rely implicitly on the invocation of the Born quantum measurement rule, whereas here decoherence derives from a density matrix formalism, obtained from employing, in place of equation 8.3.3, a functional integral representation when the QSD terms are retained. As residual SRN effects, these QSD terms lead to the random behaviour of the Born quantum measurement postulate but here the behaviour arises explicitly from *Process Physics* rather than from some meta-rule. That is, decoherence arises naturally in *Process Physics* from the limitations of self-referencing[6].

In the foregoing, there are two key aspects:

[5]Identification of the forms of the Linblad operators is beyond the scope of the present work.

[6]Self-referencing blurs the artificial exophysical distinction between object and environment; it is of at least passing interest that, for example, Zurek [369, 370, 371] has suggested a derivation of Born's rule that is based on the inclusion of the environment.

- there is a deterministic and unitary evolution that tracks and preserves topologically encoded information; and
- there are stochastic QSD terms, the form of which protects the topologically encoded information during localization events and, in QHFT, matches process time to real time, providing an ordering of events, an intrinsic temporal direction or 'arrow of time', and a modelling of the contingent present moment effect.

Emergent classicality

As the schematic in figure 8.1 illustrates, the topologically encoded information may have more than one 'footprint' in the process space and, in the induced approximate standard quantum theory, these correspond to superpositions $\psi_1(x) + \psi_2(x)$ representing quantum non-locality such as that found, for example, with the two-slit experiment for photons and also by EPR entanglement. The localization of such

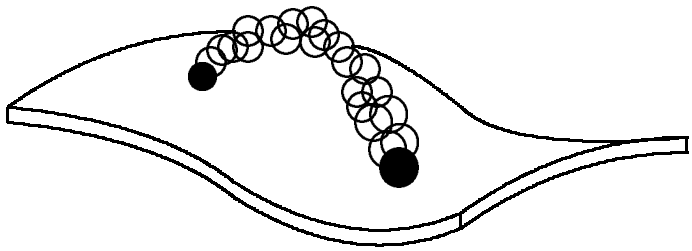


Figure 8.1: Non-local embedding of topologically encoded information illustrating two 'footprints' in the process space (drawn as 2-D with thickness and modulo gebit structure).

states is caused by the QSD terms acting non-locally – *Process Physics* generates a theory of quantum measurement because it generates both the 'objectification' associated with classical apparatus and the actual process of (partial) wave-functional

collapse as the quantum modes interact with macroscopic detectors, which are themselves permanently localized by QSD effects. Indeed, many aspects of the quantum measurement problem are resolved if one recognizes that the collapse is actively provoked by the measuring instrument itself and this occurs because the QSD process is most active when the system deviates strongly from its dominant mode, which is the ongoing relaxation to a 3-D process space, with matter surviving only because of its topological form [7]. This non-local collapse mechanism is equivalent to a continual sharpening of the homotopic mappings towards what one might consider to be a 'classical' configuration, providing what is ordinarily recognized as 3-D space.

The mappings $\pi_{\alpha\beta}$ are related to group manifold parameter spaces, with the group determined by the dynamical stability of the mappings, and gauge symmetry leads to the flavour symmetry of the standard model. Quantum homotopic mappings behave as fermionic or bosonic modes for appropriate winding numbers, thus *Process Physics* predicts both fermionic and bosonic quantum modes – not as 'objects' or 'particles' but as relational topologically encoded information. Moreover, QHFT has fractal embedded fermionic/bosonic modes and is thus much richer than conventional quantum field theory. If one introduces a meta-colour dynamics, the solution of the functional Schrödinger equation modulo QSD terms can be expressed as functional integrals that may be deconstructed, using the techniques of functional integral calculus, down to fermionic functional integrals for preons (with colour necessarily 'confined', as there are no topological defects corresponding to the preons).

[7] In essence, this matches the suggestion by Penrose [372] that quantum measurement is a manifestation of quantum gravity. Demonstrating the validity of that view, however, could only eventuate when quantum gravity is shown to derive from deeper considerations rather than *ad hoc* arguments such as quantization of classical spacetime.

8.4 Quantum Field Theory

QHFT is a highly complex accounting system for emergent properties of the stochastic network that, at an appropriate scale, should be amenable to approximation by a more familiar conventional quantum field theory (QFT), relating the deeper theory to the standard model of particle physics.

An effective QHFT should reproduce the emergence of the process space aspect of the quantum foam; the QSD processes have a key rôle in this by tending to enhance classicality, leading to the emergence of a wave-functional system $\Psi[U(x);t]$ where the configuration space is that of homotopies mapping a 3-space to $U(x) \in G$, where G is some group manifold space that describes 'flavour' degrees of freedom [18]. Thus the Schrödinger wave-functional equation for this QFT will have the form

$$\Psi[U;t+\Delta t] = \Psi[U;t] - iH\Psi[U;t]\Delta t + \text{QSD terms}, \tag{8.4.1}$$

where the general form of H is known and where the QSD terms appear as a new residual instantiation of the SRN. This system describes Skyrmions embedded in a continuum spacetime and it is significant that such Skyrmions are stable (at least, in flat space and for static Skyrmions) only if that space is 3-D.

8.4.1 Emergent 'flavour' and hidden 'colour'

As before, modulo the QSD terms (8.4.1) may be expressed by the functional integral

$$\Psi[U;t'] = \int \mathcal{D}\tilde{U}\ \exp(iS[\tilde{U}])\Psi[U;t]. \tag{8.4.2}$$

In a sense, *Process Physics* theory now turns full circle because to make sense of, and gain insight into, the phenomenology contained in equation 8.4.1 or 8.4.2, it is

convenient to recognize that functional integrals having this Skyrmionic form may be written in terms of functional integrals employing Grassmann variables, although this requires the introduction of an extra degree of freedom as a fictitious 'meta-colour' and associated fictitious coloured vector bosons – essentially, this is the reverse of the Functional Integral Calculus hadronization technique employed in the GCM of QCD ([345, 346, 373] and see §6.3 on page 153). For flat spacetime, the action for the Grassmann and vector boson part of the system has the form

$$S[\overline{p}, p, A^a_\mu] = \int d^4x \left(\overline{p}\gamma^\mu (i\partial_\mu + g\frac{\lambda^a}{2}A^a_\mu)p - \frac{1}{4}F^a_{\mu\nu}(A)F^{a\mu\nu}(A) \right), \tag{8.4.3}$$

where the Grassmann variables $p_{fc}(x)$ and $\overline{p}_{fc}(x)$ have flavour and meta-colour labels. The Skyrmions are then the low energy Nambu-Goldstone modes of this Grassmann system (other emergent modes are of higher energy and may be neglected). These colored and flavoured but fictitious fermionic fields, $\overline{p}$ and p, correspond to the proposed preon system [374, 375]. As they are purely fictitious, in the sense that they have no corresponding excitations in the system, the meta-colour degree of freedom must be hidden or confined. At this point, *Process Physics* manifests the general feature of the standard model: the existence of 'particles' with flavour and confined colour degrees of freedom, with the advantage now that with this preon form the states of the system may be more easily determined by using the more familiar language of fermions and bosons rather than working with the Skyrmionic system, with the proviso that only colour singlet states are finally permitted.[8]

[8] Consistent with remarks in the introduction to this chapter, it is important to note that equation 8.4.3 and the action in equation 8.4.2 are certainly not the final forms and further analysis will be required to fully extract the induced actions for the emergent QFT.

8.5 Gravity

8.5.1 Expansion and the Cosmological Constant

The principal coarse-grained large scale feature of Process Physics is a growing three-dimensional process space with the characteristics of the hypersphere S^3 (predictively, one of the forms permitted by Einstein's model) embedded in a random background with no metric structure. This is essentially equivalent to the 'giant component' of random graph theory demonstrated in the previous chapters and such a structure is expected to exhibit growth, as previously disconnected gebits are acquired, and decay as gebits lose their 'stickiness'. Such growth/decay behaviour is described, to first order, by the simple rate equation

$$\frac{dn}{dt} = an(t) - bn(t), \quad a \geq b \geq 0 \tag{8.5.1}$$

where the parameters a and b (most certainly dynamic variables with constraints in any realistic implementation) respectively determine the growth and decay rates and $n(t)$ is the size of the hypersphere as measured by its gebit number density. The first term arises because, ignoring the topological defects associated with quantum 'matter', success in randomly attaching new gebits is proportional to the number of gebits present (the 'stickiness' effect). Topological defects, having different stickiness, introduce second-order effects and also affect the decay rate, slowing the rate of expansion. The activity of the binder mechanism, the resource of SRN, an inexhaustible supply of disconnected gebits, and the existence of the hypersphere all ensure that $a \geq b$. Thus n(t) will show exponential growth and so *Process Physics* predicts a positive cosmological constant, consistent with recent observations of far distant supernova counts [376, 377]. Moreover, this is unrelated to the phenomenon of gravity but is to

be interpreted as deriving from the order/disorder interplay intrinsic to the dynamics of the process system.

8.5.2 Gravity, the Equivalence Principle, and Black Holes

That which is ordinarily called 'matter' is, in *Process Physics*, synonymous with topological defects embedded in the process space and such defects are expected to have a higher connectivity than the substratum – indeed, as was indicated in HPS theory, they distort the S^3 signature of the space. For this reason, fields were introduced to emulate this extra non-spatial connectivity. Now, because of the extra connectivity, gebits in the region of the matter fields are less sticky and so decay at a faster rate; they are more transient and also less effective in attracting new gebits.

This suggests an interpretation whereby matter-occupying regions of the process space act as nett sinks for gebits so that in the neighbourhood of such sinks, the process space will undergo a diffusion/relaxation procedure in which space, in effect, moves towards the matter – matter acts as a sink for space (but never as a source). That is, the diffusion/relaxation behaviour induced by the presence of matter produces an *in-fall* or *flow* effect that in turn acts on all surrounding matter. The phenomenology of this in-fall effect corresponds to, and explains, that which is called *gravity*, and the independence of the in-fall from the nature of the surrounding matter provides the analogue of the Equivalence Principle. Further, if the flow rate exceeds the rate at which 'motion' with respect to the process space is possible, an event horizon must result, leading to the formation of an inescapable sink – the black hole scenario. A logical conjecture is that a black hole singularity is comprised of a tangle of topological defects with effective dimensionality so high that the process space is

folded in upon itself, maintaining the in-flow of quantum foam. Since the in-flowing 'material' is nothing more than topologically encoded information, the in-flow represents a loss of information and so gravity itself is seen to be an information loss phenomenon.

8.6 Motion, inertia, and relativistic effects

In the 3-D process space of topologically encoded information, persistent patterns at all fractal levels abide by virtue of ongoing self-replication and it was suggested in §8.2 on page 295 that patterns may "drift through the system". This was not to suggest motion but, rather, that the patterns are independent of their specific components – there is no mechanism in *Process Physics* for a fixed pattern (corresponding, say, to a topological defect) to 'slide' *through* the quantum foam. So what could it mean for a topological defect, or indeed a vast bound collection of defects, to be 'in motion'?

Recall that the discussion concerning the rôle and attributes of the binder (§7.3.2 on page 243) indicated that the most common situation whereby active nodes arise is the case where a parent node to leaf nodes has degree greater than one (it matters not, of course, whether the nodes are taken to be monads or gebits). This is important because it means that most active nodes are located peripherally rather than being situated 'internally' and so the newest (and hence the stickiest[9]) gebits in a topological defect tend to be peripheral. Statistically, active nodes are randomly distributed so the self-replication process is statistically homogeneous and the defect may be considered to be stationary relative to the (background) quantum foam in which it

[9]The greatest probability for the topology of new gebits is that they be minimally connected i.e. tree- or augmented-tree structures with many active nodes.

is embedded. Suppose now that a situation arises – some 'event' – to disrupt this homogeneous self-replication – say, as a result of interaction with another defect, or simply as a result of local random deviations – such that gebits are preferentially replaced on one 'side'. Older gebits on the opposing 'side', in the meantime, continue to lose their stickiness, becoming still less active and decaying. The nett result is a translation, with respect to the process space, in the direction of the preferential self-replication. That is, the defect structure undergoes relative *motion* as a result of a simple self-consistency argument. Moreover, once begun, the procedure is self-perpetuating and must necessarily continue until the topological defect is subject to a new event. That is, uniform motion – *inertia* – is persistent self-replication of gebit connectivity patterns in a set direction, behaviour that is consistent with the iterator mechanism and essentially no different to the means by which stable emergent forms arise in the first place.

A further conjecture is the expectation in *Process Physics* that such motion also yields time-dilation effects because self-replication of the defect structure has two components – self-replication associated with any internal oscillations and self-replication strictly associated with motion – resulting in 'competition for resources' that slows down the internal processing, a notion discussed by Toffoli [378]. There is a long history of such effects in various non-relativistic continuum systems (see, for example Unzicker [379] for an interesting account, including a collection of ideas published by various authors over many years) and, in particular, emergent Lorentz relativistic effects have been observed in the modelling of dislocations in crystals, where so-called 'breather modes' arise as solitons [380].

The moral seems to be that the occurrence of Lorentz relativistic effects is a

generic feature in the unification of substratum systems and their embedded excitations – here, the embedded excitations are solitons, seen as a dynamical emergent feature within some dynamical background rather than as a construct 'pasted' onto some structureless *a priori* geometrical background. This dynamical interpretation of Lorentz symmetry was argued by Bell [94] and, indeed, by Lorentz. More recently, similar ideas have appeared in the analysis of sound waves in classical fluids [381, 382] and it appears reasonable to assume that QHFT also will display the time dilation and length contraction effects of generic Lorentzian dynamics.

8.7 Concluding remarks for Process Physics

Fittingly, *Process Physics* is both an end and a beginning. The commentary in this chapter marks the culmination of the present work while concurrently signalling the genesis of a major development in the modelling of reality. Proceeding from the early insights from HQS theory, *Process Physics* obtains from the introduction, via HPS theory, of non-geometric time as *process* by way of a stochastic, self-referential information theoretic bootstrap system, and has been shown capable of generating a rich phenomenology, including:

1. the emergence of primitive geometric structures (*gebits*) bearing the signature of the three-dimensional hypersphere, S^3

2. the self-linking of these gebits so as to form a fractal processing space with embedded topological defects driven by self-referential noise

3. the appearance of persistent topological defects within the substratum space interpreted, via the development of an appropriate encoding formalism, as emergent quantum 'matter' displaying flavour and hidden colour degrees of freedom and with spin 1/2 formalism also emergent

4. the self-organisation of this topologically encoded information so as to manifest quantum foam behaviour involving both space and quantum matter embedded in that space: that is, emergent quantum behaviour that derives as a logical necessity of the information theoretic stochastic process

5. fundamental non-locality of the quantum matter and of the substratum space yielding an explanation for the collapse mechanism of measurements on quantum systems, whereby the collapse is seen to be a manifestation of SRN

6. the on-going non-local collapse of that quantum structure, yielding a higher level effective classical 3-D space that undergoes expansion consistent with the observed behaviour of the universe

7. the emergence of classical behaviour, where reality at the macroscopic level appears discrete and 'hard' in contrast to the 'soft' and 'fuzzy' underlying fractal substratum of relational information

8. unification of space and the quantum leading to an explanation of gravity (displaying the equivalence principle), where matter serves as an effective sink for the information patterns that provide spatial degrees of freedom of the quantum foam, and gravity is then seen as a nett in-flow of space into matter.

9. motion, inertia, and Lorentzian relativistic effects of time dilation and length

contraction, appearing as intrinsic emergent features of the *Process Physics* schema

The unification[10] of gravity and the quantum theory that has so long eluded physics, and the explanation of classicality, motion, and inertia – similarly elusive concepts – are made possible here only by abandoning long held 'fundamental' axioms and assumptions in favour of a minimal stochastic information process system in which the fundamental properties must emerge from the system without prescription. That they appear to do so is remarkable. Certainly, these are big claims and it is acknowledged that the ideas put forward in this chapter are more suggestive than definitive – rigorous development will likely require considerable future research effort. Nonetheless, the key concepts have been outlined and they appear to point the way to a deeper, even profound, understanding of physics as it is presently 'known' and of Nature itself, which doubtless has considerably more secrets yet to reveal.

[10]Strictly, in *Process Physics*, gravity and the quantum are not unified in the usual sense – that objective is rendered redundant by the deeper description that obviates the need to force those two distinct *measurement* theories into an unhappy union.

Summary and Conclusion

Arising from dissatisfaction with the long-standing impasse confronting the physics community, this thesis has described the development of a very different view of physics at the deepest level, concluding with the presentation of the fundamentals of the theory of *Process Physics*, a radical alternative in the modelling of reality.

The work began with a review, presented in Part I, of the historical development of physics from the time of the Ancient Greeks to the present day, with particular emphasis given to the philosophic and metaphysical foundations embodied by schools of thought and specific developments that led to the mathematization of physics and the prevailing geometrical continuum model of space-time. As a result, it has been shown that this conventional, 'mainstream' approach differs little in its conception to the nâive 'container plus hard atom' model of Antiquity and, moreover, that the dominance enjoyed by this approach to modern physics has arisen more by historical accident than overt scientific design.

The subsequent examination of a range of current research activities having implications relevant both to physics at the deepest level, generally, and to space-time, more specifically, identified distinct, unequivocal, and intractably problematic flaws and limitations in the paradigms of the geometric continuum model and standard

formulations. This, then, strongly suggested that the prevailing models are not fundamental, thus demonstrating the validity of pursuing a radically different approach. Indeed, many eminent researchers share this perception and have turned their attention to devising alternative formulations (particularly in quantum gravity and quantum cosmology) the most notable of which involve some variety of *pregeometry.*

Insights gained from contemplating the disparate views of GR and quantum theory and their resonance with the confrontation over the duality of description between the static view of classical dynamics and the evolutionary view associated with entropy, served to highlight the essentially discrete nature of plausible pregeometric notions. Sharing Rosen's view that contemporary physics treats of simple systems whereas natural systems are invariably anything *but* simple, recognition of the combined aspects of discreteness and the nature of problems associated with complex systems led to a consideration of the essential features and phenomenology associated with the science of complexity. From this, it was concluded that the Universe is the quintessential autopoietic complex adaptive system, the vitally self-referential nature of which dictates the primacy of stochasticity and randomness, apparently echoed by the crisis in mathematics first identified by Gödel.

To comprehend the implications of this deep insight, the work of Gödel, Turing, and Chaitin was investigated. The inevitability of self-referencing in sufficiently rich systems was found to be intimately tied to the distinction between syntax and semantics, between exophysics and endophysics. Moreover, it was noted that randomness, or noise, follows as a direct consequence of self-referencing in accordance with the degree to which self-referencing is complete. This observation provided the key to discovering the concept of *self-referential noise* (SRN), which puts the *process* into

Process Physics, giving rise to the completely self-referential, self-modifying entity that Kampis refers to as a "Leibnizian monad" [335] – the seed exposed by breaking the cycle hierarchy of this essentially bootstrapped system to peer within, thus to see that, indeed, in the succinct and very apt words of Manthey, "outside is as inside" [270].

Proceeding from these contemplations, the various threads were drawn together to weave a cloth of a different kind, a prescription for *Process Physics* as a pregeometric theory describing a minimal, autopoietic complex adaptive system that is necessarily discrete, relational, stochastic, self-referential, self-modifying, and devoid of *a priori* constructs. To overcome the limitations of logic inherent in traditional approaches, the distinction was drawn between syntactical information systems – the language of mathematical physics – and the autopoietic semantic information systems that capture the fundamental complementarity between function and structure. The concept of a Gödelian boundary was introduced in the description of the means by which a self-referential syntactical information system could be *bootstrapped* to yield an emergent induced formal system, corresponding to conventional syntactical modelling, and the associated accessible ensemble truths of quantum statistical measurement protocols. In this, it was indicated how self-organized criticality (SOC) would act as a filter for the seeding syntax of the autopoietic SRN process.

In Part II, that prescription was put into practice by first adapting current Quantum Chromodynamics (QCD) technology to provide a foundation for modelling the developing new theory. The initial formulation, dubbed Heraclitean quantum systems (HQS), employed Grassmann algebra to construct a precursor theory that was free of classical structures or concepts in its axioms. It was shown that the algebraic

form of the theory had a corresponding representation in random graph theory and, in particular, coarse-graining the relational information between monads resulted in the most probable random graphs having connectivity that revealed an intrinsic hyperspherical signature; that is, the most probable forms emerging from a HQS were shown to have characteristics of the three-dimensional hypersphere, S^3. These forms were dubbed 'gebits' (geometrical bits).

For all that HQS theory displayed some very interesting properties, it could not produce the phenomenon of time. That failure was found to lie in the fallacy of having assumed time to be an emergent phenomenon. Recognizing, then, the truly fundamental nature of time, provided the key to further progress in the development of the successor theory of Heraclitean process systems (HPS), which recognized the iterative character needed to model reality as a self-referential, self-organizing, self-modifying information system, where the meaning of information is entirely endophysical and notions of both information and structure are purely relational constructs. HPS theory gained time-like phenomenology and *vitae essentia* with the introduction of process and the stochasticity of self-referential noise (SRN) by which the creative and spontaneous generation of spatial structure became possible, generating the 'present moment effect' customarily associated with experiential time. HPS theory demonstrated how the nature of space and the nature of time are intimately related to the 'logic of the limitations of logic'.

Finally, HPS theory was shown to extend to the theory of *Process Physics* by an introduction of the key concepts whereby *Process Physics* connects with and motivates the higher level descriptions of the standard formulations of particle physics and gravity. Further, the potential for the new theory to go beyond the old was illustrated.

Recall, from Chapter 7 (page 286), the noted parallels between HPS theory and natural hermeneutics. In the context, now, of *Process Physics*, Kampis's further generalization in his *'molecular hermeneutics'* hypothesis provides a prophetic insight when extended beyond the biological framework for which it was originally conceived – in [362], he wrote:

> This metaphysics ['molecular hermeneutics'] grasps the essence of the few examples we have presented so far: it suggests that structure is subordinated to function (while not forgetting that there are interesting aspects of structure). We propose these statements:
>
> - 'Properties' of biological compounds are *a posteriori* and relational rather than *a priori* or structural. Structures can be used in indefinitely many ways; in particular, old structures can be reused in different new functions.
> - Information content is evoked by the embedding context, and does not exist before or after.
> - Instead of isolated components, the units of analysis are situation-dependent chunks of co-existing components.
>
> In other words, instead of assuming built-in properties, we postulate interactions that bring forth what in a permanent context would be thought of as a property. This has far-reaching consequences for the entire conceptualization.

By reading, say, 'of objects' rather than "of biological compounds" or, better, striking those words altogether, the consequences are indeed "far-reaching" – in such

manner, recognizing the priority of process, the old paradigms describing a deterministic Universe governed by dynamical laws are supplanted by a new paradigm: that of a 'Biological Universe', revealed as *Process Physics* bootstraps reality from the limitations of logic.

==//==