The Diffusion of Euclidean Shape

Wilfrid S. Kendall

1. Introduction

This paper is a preliminary report on the results of an investigation into the diffusion of Euclidean shape, using computer algebra to reduce complicated intermediate calculations to an informative final form. The computer algebra takes the form of an extension to the *symbolic Itô calculus* described in W.S. Kendall (1988). A substantially more detailed treatment (including details of how to obtain the results below and a description and discussion of the necessary extensions to symbolic Itô calculus) will be provided in a later paper. The results are new and will be of interest to workers in the field of statistics of shape, and perhaps also to mathematical physicists. They provide a reinforcement of the view expressed in W.S. Kendall (1988), and further argued in my contribution to the discussion of D.G. Kendall (1989), that computer algebra and the associated equipment now form a powerful tool for probabilists and statisticians, as indeed for the mathematical scientist in general.

Suppose k particles $X_1, \ldots, X_k$ diffuse in Euclidean n-space $\mathbb{R}^n$ according to independent copies of an Ornstein-Uhlenbeck process. Thus $X_1, \ldots, X_k$ obey the system (1.1) of stochastic differential equations

$$d_I X_i = d_I B_i - \frac{\kappa}{2} X_i \, dt \quad \text{for } i = 1, \ldots, k \tag{1.1}$$

in which $B_1, \ldots, B_k$ are independent Brownian motions in $\mathbb{R}^n$ and κ is a non-negative constant, the *Ornstein-Uhlenbeck parameter*. Here and in the following we use the stochastic calculus, so $d_I X_i$ is the *Itô stochastic differential* of the random process X_i. See Rogers and Williams (1987) for an exposition, and also W.S. Kendall (1987, Section 1), for an introduction to the notation used below and some relevant geometric considerations.

Following D.G. Kendall (1977) one may consider the *(Euclidean) shape* formed by the configuration of the k diffusing particles. That is to say, one considers the stochastic evolution of those aspects of the configuration which have nothing to do with its location, orientation, or size. The resulting *diffusion of shape* has a fascinating and beautiful structure, despite the simplicity of the underlying stochastic differential system (1.1). (Strictly speaking the shape performs a diffusion only up to a random time-change; see the comment before equations (2.2), (2.3), and (2.4).) We shall describe this structure by analyzing stochastic differential systems for the stochastic evolution of collections of *shape statistics* — configuration functions depending only on the shape of the configuration in question. These collections of shape statistics will form coordinate systems for the shape diffusion.

Here is a brief summary of some statistical and probabilistic aspects of shape relevant to this paper, and a summary of previous results on shape diffusion. Recall that two configurations each of k points are said to have the same shape if one configuration can be transformed into the other by application of a sequence of translations, rotations, and dilatations. (To avoid degeneracy we stipulate that $k > 2$ and that neither configuration is composed of totally coincident points.) This conception of shape arose from a statistical problem in archaeology (Broadbent 1980; D.G. Kendall and W.S. Kendall 1980) and has been considerably developed over the last decade (see D.G. Kendall 1984, 1986, the reviews of Small 1988, and D.G. Kendall 1989, and the introductory treatment in Chapter 8 of Stoyan et al. 1987). In particular it has been established that the space Σ_n^k of k points in n-space carries a metric which is natural from statistical and probabilistic points of view. The shape spaces Σ_1^k are metrically spheres while Σ_2^k are metrically complex projective spaces. For $n \geq 3$ the shape space Σ_n^3 is metrically a hemisphere. The general shape space does not have such simple geometry and indeed if $k > n + 1$ and $n \geq 3$ then Σ_n^k is not a smooth manifold (D.G. Kendall 1989).

In the case $k = 3$ and $n = 2$ the shape space is a complex projective space of one complex dimension and is therefore isometric to a 2-sphere. The shape of the diffusing triad $X_1X_2X_3$ is actually Brownian motion on this 2-sphere, up to a random time change. This beautiful result (linked to properties of the Hopf fibration) is due to D.G. Kendall (1977) in the case when the Ornstein-Uhlenbeck parameter κ is zero. He also identified the shape diffusion in the cases of Σ_1^4 and Σ_3^3. In W.S. Kendall (1988) an implementation of stochastic calculus in the REDUCE computer algebra language (the symbolic Itô calculus mentioned above) and a description of shape in terms of *homogeneous shape coordinates* were used to identify the shape diffusion of Σ_n^3 for $n \geq 3$ and non-negative κ. Carrying the computer algebra approach further had to await a way of handling vectors of general symbolic dimension n in REDUCE . This has now been developed, and the results described below are the first fruits of this extension.

It should be noted that Carne (1988) has also made a successful study of the shape diffusion, as part of a wider study of the geometry of shape. The explicit calculations given here complement his more algebro-geometric approach. Indeed a direct connection between diffusion theory and Riemannian geometry (as described in Chapter 5 of Ikeda and Watanabe 1981) means one can deduce the shape geometry from knowledge of the shape diffusion. The paper promised above will consider connections both with Carne's results and with unpublished work of D.G. Kendall (manuscript). Carne (1988) and also Le (1988) have obtained the form of a generalized shape diffusion for points on the sphere; in this case it is no longer possible to separate shape from size so the generalization is actually a shape-and-size diffusion.

The contents of the rest of the paper are as follows. Section 2 summarizes the results describing the general shape diffusion in terms of stochastic differential systems. Expressions corresponding to three different coordinate systems are provided: the so-called homogeneous shape coordinates of normalized square side lengths, the coordinates of standardized inner products, and the coordinates corresponding to a singular values decomposition. Section 3 discusses some of the more basic questions concerning the last of these coordinate systems, which provides the most insight of the three into the behaviour of the shape diffusion. The paper concludes with Section 4, which comprises a brief discussion of topics for further work.

I am grateful to T.K. Carne, S.D. Jacka, D.G. Kendall, and Le H.L. for their helpful comments on preliminary and draft versions of this work.

2. Stochastic Differential Systems for Shape

Suppose k particles $X_1, \ldots, X_k$ diffuse in Euclidean n-space $\mathbb{R}^n$ as specified by the stochastic differential system (1.1). As noted above, we stipulate $k > 2$. Consider the *(modified) shape* $\sigma \in \tilde{\Sigma}^k_n$ of the k-tuple $\{X_1, \ldots, X_k\}$. The modified shape is defined using the enlarged symmetry group of rotations, translations, dilatations, *and* reflections (hence the notation $\tilde{\Sigma}^k_n$ rather than Σ^k_n). (Section 3 explains how to carry results over to the full shape space Σ^k_n.) Adapting W.S. Kendall (1988), the (modified) shape σ is usefully parametrized by the *homogeneous shape coordinates* given by the $\binom{k}{2}$ normalized squared side lengths of the k-tuple $\{X_1, \ldots, X_k\}$. The normalization is obtained by dividing by the *size* Σ given in (2.1):

$$\Sigma = \frac{1}{2k} \sum_i \sum_j \|X_i - X_j\|^2. \tag{2.1}$$

Thus the homogeneous shape coordinates are given by $\sigma_{ij} = \|X_i - X_j\|^2/\Sigma$ (so $\sigma_{ii} = 0$ and $\sigma_{ij} = \sigma_{ji}$). They determine the modified shape of the multiplet $\{X_1, \ldots, X_k\}$. Consider the stochastic differential system governing the evolution of Σ and σ. Using a time-change $d\tau = dt/\Sigma$ the system can be summarized by the stochastic differential equations at (2.2), (2.3), and (2.4) below. Note from (2.3) that Σ and σ are infinitesimally uncorrelated; indeed the evolution of σ in the τ time-scale is independent of Σ and the trajectory of Σ forms a sufficient (functional) statistic for the Ornstein-Uhlenbeck parameter κ. A consequence of these observations is that shape σ and size Σ form a *skew-product decomposition* of the process of 'shape-and-size' of the multiplet $\{X_1, \ldots, X_k\}$. In particular the time-change based on size Σ turns the shape σ into a genuine diffusion governed by (2.4) below.

$$\text{Drift}\left(d_I\Sigma\right) = \left\{(k-1)n - \kappa\,\Sigma\right\} dt \tag{2.2a}$$

$$(d_I\Sigma)^2 = 4\,\Sigma\, dt \tag{2.2b}$$

$$\left(d_I\sigma_{ij}\right)\left(d_I\Sigma\right) = 0 \tag{2.3}$$

$$\text{Drift}\left(d_I\sigma_{ij}\right) = \left\{2 - (k-1)\sigma_{ij}\right\} n\, d\tau \tag{2.4a}$$

$$\begin{aligned}(d_I\sigma_{ij})(d_I\sigma_{uv}) = {} & 2\left(\delta_{iu} - \delta_{iv} + \delta_{jv} - \delta_{ju}\right)\left(\sigma_{iv} - \sigma_{iu} + \sigma_{ju} - \sigma_{jv}\right) d\tau \\ & - 4\sigma_{ij}\sigma_{uv}\, d\tau.\end{aligned} \tag{2.4b}$$

Here δ_{ij} is the Kronecker symbol, equal to unity if $i = j$ but otherwise zero. Notation such as Drift $(d_I\Sigma)$ refers to the mean forward infinitesimal increment at a fixed time of Σ, where the mean is the conditional expectation given the σ-field of events determined at the fixed time. The system (2.4) is the stochastic differential system for shape diffusion in *homogeneous shape coordinates.*

Computer algebra proved convenient in finding the above formulae, although their derivation by hand is a straightforward exercise. Indeed all the formulae for stochastic differential systems in this section can be (and have been) checked manually using Itô's lemma and (somewhat laborious) formula manipulation. Finding such formulae for the first time is of course rather harder work. It is in the exploratory phase that the benefits of the computer algebra of symbolic Itô calculus really pay off.

As in the derivation of the other stochastic differential systems described below, the computer algebra procedure for deriving (2.2), (2.3), (2.4) followed closely the method expounded in W.S. Kendall (1988). REDUCE was used in its interactive mode to define expressions for Σ and σ in terms of the Ornstein-Uhlenbeck processes $X_1, \ldots, X_k$. The procedure **d** of symbolic Itô calculus was then applied to derive expressions for $d_I\Sigma$, $d_I\sigma$ in terms of $X_1, \ldots, X_k$ and $d_IX_1, \ldots, d_IX_k$. The known second-order structure of $d_IX_1, \ldots, d_IX_k$ then allowed the determination of expressions for Drift $\left(d_I\Sigma\right)$, $\left(d_I\Sigma\right)^2$, $\left(d_I\Sigma\right)\left(d_I\sigma_{ij}\right)$, Drift $\left(d_I\sigma_{ij}\right)$, and $\left(d_I\sigma_{ij}\right)\left(d_I\sigma_{uv}\right)$, in terms of $X_1, \ldots, X_k$. Finally the REDUCE package was used to determine equivalent expressions in terms of Σ, σ_{ij} as above.

The advance in technique over W.S. Kendall (1988) lies in the use of REDUCE operators representing the action of summing over dummy variables. This allows the treatment of symbolic dimension. These operators and their associated simplification rules will be described in the future paper promised above.

The results of (2.2), (2.3), (2.4) (summarizing the second-order structure of Σ, σ) formed an intermediary stage in the derivation of results concerning the standardized inner product system below, and these in turn provided a second-order structure by which was derived the system for singular values decomposition. This approach (similar to that employed in the precursor paper of Kendall, 1988) typifies a step-by-step strategy which is important in computer algebra as a means of reducing the (often extreme) length of intermediate expressions. The occurrence of machine-overflow is thereby minimized and (of equal importance) the user finds it easier to see the direction in which the interactive calculations are pointing.

From henceforth we work in the τ-timescale and consider the shape diffusion σ.

The shape-diffusion formulae at (2.4) can be re-expressed in another coordinate system determined by the inner-products of a normalized system of particles representing the shape of the multiplet. These inner-products $\{C_{ij} : i, j = 1, \ldots, k\}$ are defined by the following (in which $\overline{X} = \frac{1}{k}\sum_j X_j$):

$$C_{ij} = \langle X_i - \overline{X}, X_j - \overline{X} \rangle / \Sigma. \tag{2.5}$$

The inner-products satisfy some important relationships:

$$\sigma_{ij} = C_{ii} - 2C_{ij} + C_{jj}, \tag{2.6a}$$

$$\sum_j C_{ij} = 0 \quad \text{for all} \quad i, \tag{2.6b}$$

$$C_{ij} = \frac{1}{k}\left\{\frac{1}{2}\sum_k (\sigma_{ik} + \sigma_{jk} - \sigma_{ij}) - 1\right\}. \tag{2.6c}$$

It may be deduced from the definition of Σ that

$$\sum_i C_{ii} = 1. \tag{2.7}$$

In this new coordinate system the stochastic differential system of (2.4) transforms to the following, the stochastic differential system for shape diffusion in *standardized inner product coordinates*:

$$\text{Drift}\,\big(d_I C_{ij}\big) = \big\{\delta_{ij} - (k-1)C_{ij} - 1/k\big\}\, n\, d\tau \tag{2.8a}$$

$$\begin{aligned}(d_I C_{ij})(d_I C_{uv}) = {} & \big(\delta_{iu} - 1/k\big)C_{jv}\, d\tau + \big(\delta_{iv} - 1/k\big)C_{ju}\, d\tau \\ & + \big(\delta_{ju} - 1/k\big)C_{iv}\, d\tau + \big(\delta_{jv} - 1/k\big)C_{iu}\, d\tau \\ & - 4\, C_{ij}C_{uv}\, d\tau. \end{aligned} \tag{2.8b}$$

Formulae (2.8) were derived from (2.4) using computer algebra and then verified (once!) by hand.

Systems (2.4) and (2.8) are not particularly informative about shape diffusion (though it is worth noting that the spatial dimension n enters into the systems only through the drifts). The stochastic matrix $\mathbf{C} = \{C_{ij} : i, j = 1, \ldots, k\}$ is a symmetric non-negative definite matrix and so it is natural to consider its spectral decomposition as providing a further system of coordinates. Consider

$$\mathbf{C} = \mathbf{R}\,\Lambda\,\mathbf{R}^T \tag{2.9}$$

where $\mathbf{R} = \{R_{ij} : i, j = 1, \ldots, k\}$ is a stochastic rotation matrix formed from the eigenvectors of $\mathbf{C}$ and $\Lambda = \{\lambda_i \delta_{ij} : i, j = 1, \ldots, k\}$ is a stochastic diagonal matrix formed from the eigenvalues of $\mathbf{C}$. (The matrices $\mathbf{R}$ and Λ are related to a singular values decomposition of a standardized representation of the multiplet $\{X_1, \ldots, X_k\}$.) Suppose that $\mathbf{R}$ is defined by the *Stratonovich* stochastic differential equation

$$d_S\mathbf{R} = \mathbf{R}\; d_S\,\eta \tag{2.10}$$

where $\eta = \{\eta_{ij} : i, j = 1, \ldots, k\}$ is the *rotational noise* for the stochastic rotation process $\mathbf{R}$. The stochastic differential system for shape diffusion in *singular values decomposition coordinates* is a stochastic differential system for Λ and η such that (2.9) yields a set of standardized inner product coordinates with the correct statistics (that is to say, satisfying the standardized inner product stochastic differential system (2.8)). In fact the singular values decomposition system is not uniquely determined by this requirement if rank considerations require more than one eigenvalue to be held fixed at zero, and it transpires that the system exhibits divergence (thus failing to define completely the evolution of $\mathbf{R}$ and Λ) on collision of a pair of eigenvalues neither one being held fixed at zero. Divergence problems will also arise for $\mathbf{R}$ if a moving eigenvalue hits a couple of eigenvalues held fixed at zero. Section 3 shows that these problems do not arise in practice.

A combination of computer algebra and manual calculation (reinforced at certain points by general arguments, and verified by manual calculation) shows that apart from the above considerations the required stochastic differential system must be as follows. Note that because we centralized the configuration (thus allowing the relation (2.6b)) one of the eigenvalues is always zero. We stipulate this to be the first eigenvalue, so $\lambda_1 = 0$ for all time. Indeed considerations of the rank of $\mathbf{C}$ make it clear that a total of at least r eigenvalues must be zero at all times, where $r = \max\{k - n, 1\}$. For convenience we order the eigenvalues in ascending order and stipulate that the first r eigenvalues are to be fixed at zero, so $0 = \lambda_1 = \cdots = \lambda_r \leq$

$\lambda_{r+1} \le \cdots \le \lambda_k$. In the sequel we refer to $\lambda_{r+1}, \ldots, \lambda_k$ as the 'moving eigenvalues'. Note finally from (2.7) that $\sum_i \lambda_i = 1$.

$$\text{Drift}\left(d_I\eta_{ij}\right) = 0, \tag{2.11a}$$

$$(d_I\eta_{ij})^2 = -(d_I\eta_{ij})(d_I\eta_{ji}) = \frac{\lambda_i + \lambda_j}{(\lambda_i - \lambda_j)^2}\, d\tau \tag{2.11b}$$

when $\lambda_i \ne \lambda_j$ and neither i nor j equals unity,

$$(d_I\eta_{ij})(d_I\eta_{uv}) = 0 \tag{2.11c}$$

when the conditions for equation (2.11b) do not apply.

$$(d_I\eta_{ij})(d_I\lambda_u) = 0. \tag{2.12}$$

$$\text{Drift}\left(d_I\lambda_i\right) = -\Big[\Big\{2\Big(\sum_{j:j\ne i} \frac{1}{\lambda_j - \lambda_i}\Big) + (k-1)n\Big\}\lambda_i + k - n\Big]\, d\tau \tag{2.13a}$$

when $i > r$ (this drift is zero otherwise),

$$(d_I\lambda_i)^2 = 4\lambda_i(1-\lambda_i)\, d\tau, \tag{2.13b}$$

$$(d_I\lambda_i)(d_I\lambda_j) = -4\lambda_i\lambda_j\, d\tau \quad \text{when } i \text{ does not equal } j. \tag{2.13c}$$

REDUCE was used to create expressions for the C_{ij} in terms of λ_a and η_{uv}. Symbolic Itô calculus was then employed to find equations for the second order statistics of Λ, η in the particular case $r = 1$. Hence were derived the equations (2.11), (2.12), (2.13). The case of general r could then have been derived from a limiting argument or by modifying the computer algebra manipulations used for $r = 1$. In actual fact the correctness of the stochastic differential system for all r was then checked manually. Thus interactive computer algebra found the form of the solution, which was then verified manually to hold in all cases.

Equations (2.11), (2.12), and (2.13) reveal the structure of the shape diffusion to be that of a *skew-product* in the terminology of Pauvels and Rogers (1988), although it does not quite fall within the scope of the theory described there (since the shape diffusion will not in general be a Riemannian Brownian motion). The skew-product property follows by noting that the stochastic differential system (2.13) for Λ is autonomous and (by (2.12)) infinitesimally uncorrelated with the evolution of $\mathbf{R}$. If on the other hand Λ is conditioned to be held fixed then the stochastic differential system for $\mathbf{R}$ is that of a fixed diffusion with parameters depending on Λ. For (2.6b) and the fact that λ_1 is fixed at zero imply that $\mathbf{U} = \mathbf{R}(0)^{-1}\mathbf{R}$ leaves fixed the unit vector $(1, 0, \ldots, 0)^T$ and so under the Λ-conditioning $\mathbf{U}$ is a left-invariant diffusion on the corresponding subgroup $\mathrm{SO}(k-1) \le \mathrm{SO}(k)$. In the full-rank case $r = 1$ this conditioned diffusion is actually a Brownian motion with respect to a left-invariant Riemannian metric depending in general on the conditioned value of Λ.

If Λ is not conditioned but left free to diffuse according to (2.12) and (2.13) then in general $\mathbf{U}$ is only a Γ-martingale with respect to a left-invariant connection (see W.S. Kendall 1987 for an explanation of this term). For the left-invariant metric on $\mathrm{SO}(k-1)$ is *not* in general unique (except in the special case of $k = 3$ already covered by W.S. Kendall 1988, and the trivial case of $k = 2$ which was excluded at the outset of this paper). Consequently except in these special cases $\mathbf{R}$ cannot be expressed as a diffusion on $\mathrm{SO}(k)$ subject to a random time change controlled by Λ. Thus the decomposition is not a skew-product decomposition in this special sense (which is perhaps what is more generally understood by the term 'skew-product'). In geometrical terms the singular values decomposition does not in general decompose the Riemannian metric induced by $\mathbf{R}$ and Λ into a warped product.

3. Answers to Some Basic Questions

In this section two fundamental features of the system (2.11), (2.12), (2.13) are discussed. Only the general lines of proofs are indicated.

3.1. Whether Eigenvalues Collide

The first feature concerns whether the system for the singular values decomposition defines the shape diffusion for all time. As noted above, the stochastic differential system (2.11), (2.12), (2.13) determines the stochastic evolution of shape only up to the first time a pair of the last $k - r$ eigenvalues collide, or (only in the case $r > 1$) if λ_{r+1} hits zero. Moreover if $r > 1$ then the system is not uniquely determined by the requirement that $\mathbf{C} = \mathbf{R}\,\Lambda\,\mathbf{R}^T$ should satisfy (2.8), as one can introduce extra rotational diffusion on axes corresponding to some pairs of $\{\lambda_2, \ldots, \lambda_r\}$.

The lack of uniqueness presents no problem, since we need only to *synthesize* $\mathbf{C} = \mathbf{R}\,\Lambda\,\mathbf{R}^T$ with the correct statistics. We do not therefore require uniqueness. The system (2.11), (2.12), (2.13) must be minimal in some sense, but we will not pursue this further here.

The question of collision might present a problem. However it can be shown that if initially the 'moving eigenvalues' $\lambda_{r+1}, \ldots, \lambda_k$ are distinct then with probability one at no future time will any pair collide. This is established by considering the positive real-valued process

$$\Phi = - \sum\sum_{r+1 \le i < j \le k} \log(\lambda_j - \lambda_i). \tag{3.1}$$

The process Φ diverges to infinity precisely when a pair of 'moving eigenvalues' collide. Combination of Itô's lemma, the system (2.13), and a permutation argument for a triple sum produces an argument showing the following:

$$\text{Drift}\,\big(d_I\Phi\big) = \binom{k-r}{2}\big\{n(k-1)-2\big\}\,d\tau. \tag{3.2}$$

(A similar but more tedious argument can be applied to evaluate $\big(d\Phi\big)^2$.) This implies that the process

$$\Psi(\tau) = \Phi(\tau) - \binom{k-r}{2}\big\{n(k-1)-2\big\}\tau \tag{3.4}$$

defines a continuous local martingale. As such it may be expressed as a random time-change of real-valued Brownian motion. On any compact time-interval $[0,T]$ the trajectory of Ψ is bounded below by $-\binom{k-r}{2}\big\{n(k-1)-2\big\}T$ (by virtue of the positivity of Φ). Consequently it follows that in any given compact time-interval with probability one Ψ and hence Φ must be bounded above by random but finite bounds. (One appeals to the properties of real-valued Brownian motion.)

The above shows that Φ remains finite for all time, and so no pair of 'moving eigenvalues' may collide. We see in Subsection 3.2 that λ_{r+1} will not hit zero if $r > 1$. These arguments show that the system (2.11), (2.12), (2.13) defines the shape diffusion for all time, so long as the initial values of the last $k-r$ eigenvalues are distinct. In effect the shape diffusion is thereby defined in coordinates of the singular values decomposition over all of $\tilde{\Sigma}^k_n$ except on a polar subset

$$\mathcal{P} = \{\lambda_i = \lambda_j \text{ for some pair } i \neq j \text{ with } r < i < j \leq k\}.$$

(A polar subset is one which the shape diffusion never visits after time zero.) This is in close analogy to the way in which the classical expression of Euclidean Brownian motion in polar coordinates (using a Bessel process) breaks down at the origin.

Of course systems (2.8) or (2.4) provide definitions of the shape diffusion holding over all of $\tilde{\Sigma}^k_n$ without exception. Further investigation of the polar subset $\mathcal{P}$ would involve exploitation of the connection between Riemannian geometry and diffusion theory, alluded to in the introduction. This will be discussed in the follow-up paper promised above.

3.2. The Full Shape Diffusion

The second feature concerns the fulfilment of the promise in Section 2 to show how to derive formulae for the full shape diffusion on Σ^k_n. The answer hangs on determining precisely when λ_{r+1}, the 'smallest moving eigenvalue', can ever hit zero.

First note that $\tilde{\Sigma}^k_n = \Sigma^k_n$ in the case $k \leq n$ (since the symmetry group $\mathrm{SO}(n)$ can always carry multiplets of n or fewer points into their mirror images). Thus for $k \leq n$ the full shape diffusion is already identified. It remains to discuss the case $k > n$, which will split into the 'critical case' $k = n + 1$ and the case $k > n + 1$.

It is convenient to digress at this point to establish the behaviour of λ_{r+1} in the general case. The 'smallest moving eigenvalue' λ_{r+1} can hit zero if and only if the critical case $r = k - n = 1$ holds (which completes the argument of Subsection 3.1 to show the good behaviour of the system at (2.11), (2.12), (2.13)). This result is proved by comparing λ_{r+1} to a Bessel process of appropriate dimension.

First note that if $n = 1$ then there is nothing to prove, as $r = k - 1$, $\sum_i \lambda_i = 1$, and so $\lambda_{r+1} = \lambda_k = 1$ is constant.

Suppose $n > 1$. Let $T^{(\epsilon)}$ be the first time at which $\lambda_{r+2} - \lambda_{r+1}$ is no larger than ϵ. Assuming the 'moving eigenvalues' are initially distinct, $T^{(\epsilon)}$ is positive for sufficiently small ϵ. By the no-collision result above, $T^{(\epsilon)} \to \infty$ as ϵ tends to zero.

Consider λ_{r+1} in a new time-scale suggested by (2.13b) and defined by $d\tilde{\tau} = (1 - \lambda_{r+1})\, d\tau$. Working up to the random time $\tilde{T}^{(\epsilon)} = \int_0^{T^{(\epsilon)}} (1 - \lambda_{r+1})\, d\tau$ the evolution of λ_{r+1} in the new time-scale is governed by

$$\operatorname{Drift}\left(d_I \lambda_{r+1}\right) = -\,2H\lambda_{r+1}\, d\tilde{\tau} + \nu\, d\tilde{\tau} \tag{3.5a}$$

$$\left(d_I \lambda_{r+1}\right)^2 = 4\lambda_{r+1}\, d\tilde{\tau} \tag{3.5b}$$

where $\nu = 2r - (k - n)$ and

$$H = \left\{ \left(\sum_{b > r+1} \frac{1}{\lambda_b - \lambda_{r+1}} \right) + \frac{(k-1)n - \nu}{2} \right\} (1 - \lambda_{r+1})^{-1}. \tag{3.6}$$

Note that H is bounded over the time interval $0 < \tilde{\tau} \leq \tilde{T}^{(\epsilon)}$ since over this interval $1 - \lambda_{r+1} > \lambda_b - \lambda_{r+1} > \lambda_{r+2} - \lambda_{r+1} \geq \epsilon$.

From (3.5) it follows that the process $X = \sqrt{\lambda_{r+1}}$ is a Bessel process of dimension ν with superimposed drift $-HX\, d\tilde{\tau}$. But H is bounded up to $\tilde{T}^{(\epsilon)}$ and so the Girsanov change-of-measure theorem implies that X can hit zero if and only if a Bessel process of dimension ν (without superimposed drift) can hit zero. Now it is classical (recalling the expression of Bessel processes as the radial parts of Euclidean Brownian motions) that such hitting of zero is possible if and only if $\nu = 1$. This means the 'smallest moving eigenvalue' λ_{r+1} can hit zero if and only if the critical case $r = k - n = 1$ holds.

Suppose $k > n + 1$, so that λ_{r+1} does not hit zero. Geometrical arguments show that the obvious projection of full shape onto modified shape

$$\pi : \Sigma_n^k - \{\lambda_{r+1} \circ \pi = 0\} \to \tilde{\Sigma}_n^k - \{\lambda_{r+1} = 0\} \tag{3.7}$$

is then a two-to-one map and indeed a covering map. The modified shape diffusion stays away from $\{\lambda_{r+1} = 0\}$ in this case, so it is possible to lift the path of the modified shape diffusion via π to the full shape space Σ^k_n. Thus if the initial point of the full shape diffusion is specified then its evolution can be deduced from that of the modified shape diffusion. So the system (2.11), (2.12), (2.13) does in fact specify the full shape diffusion in the case $k > n+1$ as well as in the case $k \leq n$.

If $k = n+1$ then the operation of taking the signed volume (via a determinant) shows that $\Sigma^{k=n+1}_n - \{\lambda_2 \circ \pi = 0\}$ can be split into two components $\Sigma^{(+)}{}^{n+1}_n$ and $\Sigma^{(-)}{}^{n+1}_n$ (recall that $r+1 = 2$ in this critical case). Moreover the reflection symmetry provides an isomorphism of the full shape diffusion on one component to the full shape diffusion on the other, and the component shape diffusions are isomorphic to the (modified) shape diffusion on $\tilde{\Sigma}^{n+1}_n - \{\lambda_2 = 0\}$. A proper description of the full shape diffusion must explain how the sign of the volume alters when the random process λ_2 visits zero.

In this case the Bessel process argument above shows that λ_2 behaves as a random time change of the square of a real-valued Brownian motion, modified by a locally bounded drift. Define $Y = \pm\sqrt{\lambda_2}$, where the sign is chosen according to the sign of the signed volume of the full shape. Itô calculus and excursion theory can be applied to show

$$\text{Drift}\left(d_I Y\right) = -\left\{\left(\sum_{j:j>2} \frac{1}{\lambda_j - Y^2}\right) + \frac{n^2-1}{2}\right\} Y\, d\tau, \tag{3.8a}$$

$$\left(d_I Y\right)^2 = (1 - Y^2)\, d\tau, \tag{3.8b}$$

$$\left(d_I \lambda_i\right)\left(d_I Y\right) = -2\lambda_i Y\, d\tau \qquad \text{when } i \neq 2, \tag{3.8c}$$

$$(d_I \eta_{ij})(d_I Y) = 0. \tag{3.8d}$$

If (3.8) is used to replace the corresponding parts of (2.13), (2.12) then we obtain an expression for the full shape diffusion in the case $k = n+1$, using a variation on the coordinates of the singular values decomposition based on η, $0 = \lambda_1$, $Y = \pm\sqrt{\lambda_2}$, and $\lambda_3, \ldots, \lambda_k$.

4. Conclusion

The work above raises a number of questions.

4.1. Relationship to Geometry

As has already been noted, Carne (1988) and D.G. Kendall (manuscript) have considered the Riemannian geometry natural to the Euclidean shape space Σ^k_n. The stochastic differential systems for shape diffusion carry within themselves information about this Riemannian geometry. The shape diffusion can be expressed as Brownian motion on the corresponding manifold modified by a drift. On the other hand the Riemannian metric tensor can be identified from the information summarized in the second-order part of the stochastic differential system for the shape diffusion. For example the level sets in $\tilde{\Sigma}^k_n$ obtained by fixing a value for $\mathbf{R}$ are incomplete but otherwise totally geodesic $(k-r-1)$-dimensional submanifolds of constant positive sectional curvatures $+1$. Indeed the 'eigenvalue map' sending $\sigma \in \tilde{\Sigma}^k_n - \mathcal{P}$ to $(\lambda_{r+1}, \ldots, \lambda_k)$ is a Riemannian submersion of the non-polar part of the modified shape space onto an open fragment of a $(k-r-1)$-sphere of constant positive sectional curvatures $+1$.

An obvious objective is to construct a set of computer algebra procedures to identify various features of the Riemannian geometry from the diffusion characteristics (D.G. Kendall and Le have carried out a similar task, using computer algebra to derive formulae for curvature for various coordinatizations of shape spaces). Account will have to be taken of the need to *complete* the Riemannian geometry to extend over the polar set $\mathcal{P}$ where pairs of moving eigenvalues coincide.

The case of Σ^3_n for $n \geq 3$ is informative. In this case $\tilde{\Sigma}^3_n = \Sigma^3_n$, $r = \max\{3-n, 1\} = 1$, $\lambda_1 = 0$ and $\lambda_2 + \lambda_3 = 1$. The eigenvalues provide one degree of freedom in a coordinate space looking like $[0, \frac{1}{2}]$. In the rotational component the only variation is provided by $\eta_{23} = -\eta_{32}$ and so the rotational coordinate space looks like a circle SO(1). Thus the singular values decomposition is based on a cylinder $[0, \frac{1}{2}] \times \mathrm{SO}(1)$. The singularity set is the circle $\mathcal{P} = \{\lambda_2 = \frac{1}{2}\} \times \mathrm{SO}(1)$ and in fact the associated Riemannian geometry collapses this circle to a point, and gives Σ^3_n the geometry of a hemisphere (this corresponds to the route followed in W.S. Kendall, 1988). The identification of $\mathcal{P}$ to a point arises from the divergence as $\lambda_3 - \lambda_2$ converges to zero of

$$\begin{aligned}\left(d_I \eta_{32}\right)^2 &= \frac{\lambda_3 + \lambda_2}{\left(\lambda_3 - \lambda_2\right)^2}\, d\tau \\ &= \frac{d\tau}{\left(\lambda_3 - \lambda_2\right)^2}. \end{aligned} \tag{4.1}$$

This example is a useful prototype for the way in which the diffusion establishes the geometry; near the polar locus $\mathcal{P}$ the divergent diffusion coefficients of the rotational noise lead to identifications in the system of coordinates of singular values decomposition. Note however in general it is necessary to take account of singularities in the Riemannian geometry.

Thus the task which should be undertaken next is to provide means, using computer algebra, of passing from the diffusion to the geometry and (if possible) taking account of identifications such as above.

4.2. Wishart Matrices

If the multiplet $\{X_1, \ldots, X_k\}$ is composed of independent random points possessing the same multivariate spherically symmetric Gaussian distribution then the resulting distribution of shape corresponds to a certain Wishart distribution normalized to have unit trace. This distribution is the invariant distribution of the shape diffusion if the points of the multiplet diffuse according to (1.1). Corresponding to this, the results above could be obtained as consequences of a stochastic calculus version of Wishart distribution theory. See the work of Bru (1989).

This 'stochastic calculus of multivariate statistical analysis' will provide the next testing ground for symbolic Itô calculus. As a further prospect there is the challenging task of generalizing to the non-spherically-symmetric extension of (1.1). The work of Mardia and Dryden (1989) suggests other exercises connected to non-central Wishart distributions.

4.3. Relevance of the λ_i Coordinates to Shape Theory

The coordinates of singular values decomposition make explicit a natural $\mathrm{SO}(k-1)$ (indeed, $\mathrm{O}(k-1)$) symmetry for Σ_n^k, allowing for certain questions a reduction of dimensionality by considering only the $k-r$ 'moving eigenvalues'. For example the locus of collinear multiplets is invariant under this symmetry, and so natural measures of distance from collinearity will be given by expressions involving only $\lambda_{r+1}, \ldots, \lambda_k$. For this reason the relative simplicity of the system (2.13) is particularly satisfying, and the Riemannian submersion referred to in Subsection 4.1 is of practical importance.

4.4. Matrix Factorization in Stochastic Calculus

We have already noted similarities to the work of Pauvels and Rogers (1988). See also Norris, Rogers, and Williams (1986) and references therein to work of Dynkin, Dyson, McKean, and Orihara on random matrices. Taylor (1988) expounds work of Malliavin and Malliavin which forms a more geometric approach to similar problems for Brownian motion on symmetric spaces. However shape diffusions appear to lack too much symmetry for any of this previous work to apply directly.

4.5. Automatic Reduction of Stochastic Differential Equations

One way to view the work of this paper is as an exploitation of a not entirely evident $\mathrm{O}(k-1)$ symmetry to reduce a stochastic differential system (2.4) or (2.8) to a form (2.11), (2.12), (2.13) involving a reduction in dimensionality. In effect, a stochastic differential system has been partially 'solved'. This raises the enticing prospect of building sets of procedures in REDUCE or another computer algebra package which would search for possible symmetries in a stochastic differential system. Having found a symmetry, this would be exploited to produce a new representation of the system in the manner given above. Sets of REDUCE procedures already exist to perform similar tasks for partial differential equations, so this prospect must be eminently achievable! From this point of view the work of this paper, originally undertaken primarily to further elicit the structure of shape diffusion, becomes a test case suggesting geometric perspectives and algorithms for complex stochastic systems.

Acknowledgement

Part of this work was aided by the provision of a Sun workstation under SERC grant GR/E 39891 while the author was affiliated to Strathclyde University.

References

Ambartzumian, R.V. (1987). *Stochastic and Integral Geometry*. Reidel, Dordrecht/Boston.

Broadbent, S.R. (1980). Simulating the ley-hunter (with discussion). *Journal of the Royal Statistical Society A* 143, 109–140.

Bru, M.F. (1989). Processus de Wishart. *Comptes Rendus de l'Académie des Sciences (Paris)* 308, 29–32.

Carne, T.K. (1988). The geometry of shape spaces. Preprint.

Durrett, R. and Pinsky, M.A. (1988). *The Geometry of Random Motion*. Contemporary Mathematics 73, American Mathematical Society, Providence, R.I.

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North Holland/Kodansha, Amsterdam/Tokyo.

Kendall, D.G. (manuscript). Ten years late better than never. Unpublished manuscript, circulated Summer 1987.

——— (1977). The diffusion of shape (abstract). *Advances in Applied Probability* 9, 428–430

——— (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society* 16, 81–121.

——— (1986). Further developments and applications of the statistical theory of shape. *Teorija Verojatnostei* 31, 467–473. Translated as: *Theory of Probability* 31 (1987), 407–412.

——— (1989). A survey of the statistical theory of shape (with discussion). *Statistical Science*, to appear.

Kendall, D.G. and Kendall, W.S. (1980). Alignments in two-dimensional random sets of points. *Advances in Applied Probability* 12, 380–424.

Kendall, W.S. (1987). Stochastic differential geometry: an introduction. In Ambartzumian (1987), reprinted from *Acta Applicandae Mathematicae* 9, 29–60.

——— (1988). Symbolic computation and the diffusion of shapes of triads. *Advances in Applied Probability* 20, 775–797.

Le, H.L. (1988). Unpublished Ph.D. thesis, University of Cambridge.

Mardia, K. and Dryden, I.L. (1989). The statistical analysis of shape data. *Biometrika*, to appear.

Norris, J.R., Rogers, L.C.G., and Williams, D. (1986). Brownian motion of ellipsoids. *Transactions of the American Mathematical Society* 294, 757–765.

Pauvels, E.J. and Rogers, L.C.G. (1988). Skew-product decompositions of Brownian motions. In Durrett and Pinsky (1988, pp. 237–262).

Rogers, L.C.G. and Williams, D. (1987). *Diffusions, Markov Processes, and Martingales*, Volume 2. Wiley, Chichester.

Small, C.G. (1988). Techniques of shape analysis on sets of points. *International Statistical Review* 56, 243–257.

Stoyan, D., Kendall, W.S., and Mecke, J. (1987). *Stochastic Geometry and its Applications*. Wiley/Akademie Verlag, Chichester/Berlin.

Taylor, J.C. (1988). The Iwasawa decomposition and the limiting behaviour of Brownian motion on a symmetric space of non-compact type. In Durrett and Pinsky (1988, pp. 303–332).

Department of Statistics
University of Warwick
Coventry CV4 7AL.