Stochastic Processes in Vision: From Langevin to Beltrami

Nir A. Sochen
Department of Applied Mathematics
University of Tel-Aviv
Tel-Aviv 69978
sochen@math.tau.ac.il

Abstract

Diffusion processes which are widely used in low level vision are presented as a result of an underlying stochastic process. The short-time non-linear diffusion is interpreted as a Fokker-Planck equation which governs the evolution in time of a probability distribution for a Brownian motion on a Riemannian surface. The non linearity of the diffusion has a direct relation to the geometry of the surface. A short time kernel to the diffusion as well as generalizations are found.

1. Introduction

There is a close and deep relation between partial differential equations (PDEs) and stochastic processes. We study this relation in the context of the scale-space approach to image processing and understanding.

The relation between Brownian motion on the plane and the two-dimensional Laplacian is well known. It is directly relevant to linear scale-space. Grey-levels are thought of as particles. A given gray level at a point (x, y) on the plane is interpreted as the number of particles at this point, or alternatively the probability to find a particle at this point. The particles perform Brownian motion on the plane and the new distribution as a function of time, or the probability to find a particle in a point, obeys the linear heat equation.

We generalize this relation and think about the non-linear diffusion processes which are applied in image processing as result of an underlying Brownian motion on a Riemannian surface (or higher dimensional manifold for higher dimensional data). In this way we rediscover the Beltrami flow which was advocated recently [5, 7, 2, 8]. It is further generalized by the choice of more general stochastic process and in particular by choosing a different drift term. Another benefit from this approach is the ability to construct a short-time kernel for the non-linear diffusion equation. This enables us to relate the geometric PDEs and the non-linear filtering approach.

The methods we describe here are well known in the mathematical literature (see [1] and references therein) and in physics. We follow here J. Zinn-Justin [10] where the subject goes under the name of stochastic quantization. We bring it here in great detail with the belief that this point of view and the techniques introduced may lead to new, interesting and useful ways of image understanding and analysis.

2. The Differential Geometric framework

The Beltrami framework of image presentation and analysis is based on geometric ideas adopted from general relativity and high energy physics. The essence of the method can be summarized as follows:

- An image is considered to be a Riemannian manifold embedded in a higher dimensional Riemannian manifold which is called the spatial-feature manifold. A two-dimensional image is according to this viewpoint a Riemannian surface. We introduce on the nonlinear surface a local coordinate system (σ^1, σ^2). The embedding of this surface in, for example, a three-dimensional space with coordinates (X^1, X^2, X^3), is realized by specifying, for each point of the surface, the three-dimensional coordinates, namely:

  $$(X^1(\sigma^1, \sigma^2), X^2(\sigma^1, \sigma^2), X^3(\sigma^1, \sigma^2)).$$

  Grey-level image, for example, is represented from this view point as the map

  $$(X^i(\sigma^1, \sigma^2) = \sigma^i, X^3(\sigma^1, \sigma^2) = I(\sigma^1, \sigma^2))$$

  where $i = 1, 2$.

  Let M denote the higher dimensional spatial-feature manifold. We introduce, in general, a one-parameter family of embedded images $(X^i(\sigma^1, \sigma^2; t))_{i=1}^{\dim M}$, where t is the independent variable of the evolution, called the scale or "time". This parameter determines the degree of blurring or denoising of the image.

- From a geometrical viewpoint this family of embedded images describes a flow of a two-dimensional surface within a higher dimensional space. The dynamics of the surface flow is governed by a nonlinear heat-type partial differential equation applied to this one-parameter family of images. The equation is derived as a gradient descent of a functional that weight embedding maps in a geometric way. It gives a precise control on the direction and amount of diffusion at each point of the image surface. This is to be compared with linear scale-space that diffuses "blindly" or the Perona-Malik equation that has a local control on the amount of diffusion but not on its direction.

We turn now to a rigorous treatment of these ideas and present the technical tools implemented in the sequel. A precise definition of a manifold and its geometry are incorporated. Next a measure on the space of embedding maps is introduced. The measure, or the energy functional, depends on the geometry of the spaces involved and is independent of the coordinate system selected to describe these manifolds.

2.1. The Induced Metric

Let $\mathbf{X} : \Sigma \to M$ be an embedding of $(\Sigma, (g_{\mu\nu}))$ in $(M, (h_{ij}))$, where Σ and M are Riemannian manifolds and $(g_{\mu\nu})$ and (h_{ij}) are their metrics respectively. We can use the knowledge of the metric on M and the map $\mathbf{X}$ to construct the metric on Σ. This procedure, denoted formally by $(g_{\mu\nu})_\Sigma = \mathbf{X}^*(h_{ij})_M$ and called the *pullback*, is given explicitly as follows:

$$g_{\mu\nu}(\sigma^1, \sigma^2) = h_{ij}(\mathbf{X})\partial_\mu X^i \partial_\nu X^j, \tag{1}$$

where $i, j = 1, ..., \mathrm{dim}M$ are being summed over, using the Einstein summation convention, and $\partial_\mu X^i \equiv \partial X^i(\sigma^1, \sigma^2)/\partial\sigma^\mu$.

For the grey-level image the induced metric is

$$\begin{pmatrix} 1 + I_x^2 & I_x I_y \\ I_x I_y & 1 + I_y^2 \end{pmatrix} \tag{2}$$

2.2. The Measure On Maps

The diffusion equation to be used is derived as a gradient descent of an action functional. The functional in question depends on *both* the image manifold and the embedding space. Denote by $(\Sigma, (g_{\mu\nu}))$ the image manifold and its metric, and by $(M, (h_{ij}))$ the spatial-feature manifold and its metric. Then, the mapping $\mathbf{X} : \Sigma \to M$ is assigned, by the Polyakov action [4], the following real number:

$$S[X^i, g_{\mu\nu}, h_{ij}] = \int d^m\sigma \sqrt{g} g^{\mu\nu} \partial_\mu X^i \partial_\nu X^j h_{ij}(\mathbf{X}), \tag{3}$$

where m is the dimension of Σ, g is the determinant of the image metric and $g^{\mu\nu}$ is the inverse of the image metric. The range of indices is $\mu, \nu = 1, \ldots, \dim \Sigma$, and $i, j = 1, \ldots, \dim M$. The metric of the embedding space is h_{ij}.

Note that the volume element as well as the rest of the expression is invariant under reparameterization, that is, $\sigma^\mu \to \tilde{\sigma}^\mu(\sigma^1, \sigma^2)$. The Polyakov action depends, actually, on the geometry and not on the way we describe the objects via our parameterization of the coordinates. In other words the resultant value of the functional does not depend on the choice of local coordinates.

2.3. The Gradient Descent Flow

Given the above functional, we have to choose the minimization criterion. We may choose, for example, to minimize the functional only with respect to the embedding. In this case the metric $g_{\mu\nu}$ is treated as a set of parameters that can be selected with reference to the application. Another choice is to minimize only with respect to the feature coordinates of the embedding space, or one may choose to minimize the image metric as well. Each of these choices yields a different flow. Some flows are, in fact, identical to existing methods like the heat flow, the Perona-Malik flow, or the mean-curvature flow.

Another important point is the choice of the embedding space and its geometry. In general, we need information about the task at hand in order to fix the right geometry [8].

Using standard methods in calculus of variations the Euler-Lagrange (EL) equations, with respect to the embedding, are (see [7] for derivation):

$$-\frac{1}{2\sqrt{g}} h^{il} \frac{\delta S}{\delta X^l} = \frac{1}{\sqrt{g}} \partial_\mu(\sqrt{g} g^{\mu\nu} \partial_\nu X^i) + \Gamma^i_{jk} \partial_\mu X^j \partial_\nu X^k g^{\mu\nu}, \tag{4}$$

where Γ^i_{jk} are the Levi-Civita connection coefficients, with respect to the spatial-feature metric h_{ij} defined as

$$\Gamma^i_{jk} = \frac{1}{2} h^{il}(\partial_j h_{lk} + \partial_k h_{jl} - \partial_l h_{jk}). \tag{5}$$

The second term in Eq. (4) is due to the non-linear geometry of the embedding space.

We view the scale-space as a gradient descent:

$$X^i_t \equiv \frac{\partial X^i}{\partial t} = -\frac{1}{2\sqrt{g}} h^{il} \frac{\delta S}{\delta X^l}. \tag{6}$$

A few remarks are in order. First, note that we took the freedom to multiply the EL equations by a strictly positive function and a positive definite matrix.

This factor is the simplest one that does not change the minimization solution, while giving a reparameterization invariant expression. This choice guarantees geometric flow and does not depend on the parameterization. The operator acting on X^i in the first term of Eq. (4) is the natural generalization of the Laplacian from flat spaces to manifolds, called the Laplace-Beltrami operator, or in short *Beltrami operator*, denoted by Δ_g. When the embedding is in a Euclidean space with a Cartesian coordinate system, the connection elements are zero. If the embedding space is not Euclidean, we have to include the Levi-Civita connection term since it is not identically zero any more.

3. Stochastic process on Riemannian manifold

3.1 Linear scale-space

We show in this subsection that the diffusion equation (6) that results from the gradient descent minimization of the Polyakov action Eq. (3) can be derived and interpreted in the framework of stochastic calculus. This new interpretation enables us to generalize and propose other flows. It is also possible in this new viewpoint to derive a short time kernel that gives us a link between PDEs based denoising and non-linear filtering techniques.

we opt in this subsection to give motivation and intuition and to refer the reader to the appendices and to textbooks for the mathematical details.

Imagen we have a particle at the origin at time $t = 0$. This particle start to move randomly performing what is called Brownian motion. Its dynamics is fixed by the equation

$$\frac{\partial \vec{q}(t)}{\partial t} = \vec{\nu}(t)$$

where $\vec{q}(t)$ are the coordinates of the particle at time t and $\vec{\nu}(t)$ is a random variable drown at any time t from a multi-dimensional normal distribution:

$$d\rho(\nu(t)) = \frac{d\nu(t)}{\sqrt{2\pi\Omega}} \exp\left(-\frac{||\nu||^2}{2\Omega}\right)$$

These equations mean that our particle has at each time a bigger chance to stay where it is or move a little. It has a small probability to have a large change in its place. It is no surprise, thus, that the probability of finding the particle at point $\vec{q} = \vec{\sigma}$ at time t is the normal distribution. We also expect that the probability to find the particle far from the origin increases with time. Given that the particle starts at the origin at time $t = 0$, i.e. $p(q, t = 0) = \delta(q)$, The probability distribution is found to be a Gaussian

$$\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{||\vec{q}||^2}{2t}\right)$$

whose variance increases with time.

This Gaussian is the Kernel (or Green function) of the PDE

$$P_t = \Delta P$$

that governs the time evolution of $P(q,t)$ with the initial condition $P(q, t = 0) = P_0(q)$.

Under the identification $I(q,t) = P(q,t)$ this is exactly the linear scale-space theory for images[witkin, Koenderink]. This identification gives the image a probabilistic interpretation. We can think about the (normalized) gray-level as the probability to find a random particle given an initial probability distribution at time $t = 0$.

3.2 Random walk on a manifold

In section 2 we reviewed the Beltrami framework that put forward the idea that the non-linearity in various PDE approaches to low and middle level vision can be understood as the affect of the low level vision objects being Riemannian manifolds. The non-linearity is encoded in this approach in the Riemannian structure i.e. the metric.

we follow the same logic here and derive a non-linear diffusion for the probability distribution by generalizing the random walk process in a geometric way. We introduce a random walk on a Riemannian manifold. The probability distribution obeys the Beltrami flow.

In order to be more explicit we introduce the vielbiens. The vielbiens are the matrices that rotate the basis vectors from the cartesian coordinates in the parametric space to the coordinate basis of the tangent space of the manifold. Let $\hat{x}^i$ be a unit vector in the cartesian i direction and $\hat{\sigma}^\mu$ the unit tangent vector which is tangent to the μ coordinate curve on the manifold then

$$\hat{x}^i = e^i{}_\mu \hat{\sigma}^\mu \tag{7}$$

Clearly the vielbiens rotate the unit matrics to the manifold's metric:

$$g_{\mu\nu} = e^i{}_\mu \delta_{ij} e^i{}_\nu \quad ; \quad g^{\mu\nu} = e^\mu_i \delta^{ij} e^\nu_i \quad ; \quad \delta^\mu_\lambda = \delta^j_i e^\mu_j e^i{}_\lambda \tag{8}$$

where $g^{\mu\nu}$ is the inverse of the metric:

$$g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_\lambda, \tag{9}$$

also e^i_μ is the inverse vielbien and summation over repeated indices is assumed.

There is a family of solutions for the gray level induced metric Eq. (2). One simple choice is:

$$(e^i{}_\mu) = \frac{1}{\sqrt{1+I_x^2}} \begin{pmatrix} 1+I_x^2 & I_x I_y \\ 0 & \sqrt{1+I_x^2+I_y^2} \end{pmatrix}.$$

Suppose that we are given that the coordinates q^μ on Σ satisfy the Langevin equation:

$$\dot{q}^\mu(t) = \frac{1}{2}\, e^\mu_{\ i} \nabla_\rho e^\rho_{\ i}(q) + e^\mu_{\ i}(q)\nu_i \tag{10}$$

where $q^\mu = q^\mu(t)$ and $\dot{q}^\mu(t) = \frac{\partial}{\partial t} q^\mu(t)$. Note that we use below the Itô calculus. In the case of the Stratanovich calculus the Langevin equation has no drift term i.e.

$$\dot{q}^\mu(t) = e^\mu_{\ i}(q)\nu_i.$$

Definition 1 *The probability distribution $P(\vec{\sigma}, t)$ is defined as*

$$P(\vec{\sigma}, t) = \langle \prod_\mu \delta(q^\mu(t) - \sigma^\mu) \rangle$$

σ^μ are time independent coordinates on the manifold Σ.

The meaning of P is the probability for finding the particle, which satisfies the stocahstic Langevin equation, at point σ at time t. If, thus, our particle is located at the origin at time $t = 0$ it is clear that the probability to find him very far from the origin after a very short time is zero. In fact the highest probability is to find him at the origin and the probability drops as we go further from the origin.

We are interested in the evolution in time of this probability distribution given an initial probability distrobution at time $t = 0$.

The time evolution of the probability distribution is given by:

Theorem 1

$$\dot{P}(\vec{\sigma}, t) = \frac{1}{2}\, \partial_\nu \left(\sqrt{g} g^{\nu\mu} \partial_\mu \left(\frac{1}{\sqrt{g}} P \right) \right). \tag{11}$$

where $\dot{P}(\vec{\sigma}, t) = \partial P(\vec{\sigma}, t)/\partial t$.

We interpret the gray level intensity as a probability distribution up to a multiplicative factor. Define:

$$I(\vec{\sigma}, t) = \frac{1}{\sqrt{g(\vec{\sigma})}} P(\vec{\sigma}, t) \tag{12}$$

then

$$I_t(\vec{\sigma}, t) = \frac{1}{2\sqrt{g}} \partial_\mu (\sqrt{g} g^{\mu\nu} \partial_\nu I). \tag{13}$$

Polyakov action.

We can use a more general Langevin equation with a genuine drift term (the drift term that we have in the above computation is a compensation term needed in order to work with the Itô calculus instead of the Stratanovich calculus). The Langevin equation, in this case, is

$$\dot{q}^\mu(t) = -\frac{1}{2}\, A^\mu(q(t)) + \frac{1}{2}\, e^\mu_{\ i} \nabla_\rho e^\mu_{\ i}(q) + e^\rho_{\ i}(q)\nu_i. \tag{14}$$

The computation of the Fokker-Planck equation is similar to the one we detailed above and the result is

$$I_t(\vec{\sigma}, t) = \frac{1}{2\sqrt{g}} \partial_\mu \big(\sqrt{g} (g^{\mu\nu} \partial_\nu I + A^\mu I) \big). \tag{15}$$

The effect of this term on the image smoothing process is under study both from a theoretical and a practical point of views [**?**]. We notice that it gives an effective connection on the manifold which is different from the Levi-Civita one.

One should be aware that the differential operator depends itself on the intensity I, which makes the the operator non-linear. This means that our analysis holds for short time interval under the assumption the metric changes smoothly and in a controlled way. This non-linearity prevents the existence of a global (in time) kernel. A kernel for a short time evolution is possible to obtain. A first approximation of this kernel is derived in the next section.

4. The Short Time Kernel

We discretize the time $t - t_0 = n\epsilon$ and the Langevin equation is:

$$q^\mu(t + \epsilon) = q^\mu(t) + \frac{1}{2}\, e^\mu_{\ i} \nabla_\nu e^\nu_{\ i}(q) + e^\mu_{\ i}(q)\nu_i\,. \tag{16}$$

The measure on the noise ν is

$$d\rho(\nu) = \frac{d\nu}{\sqrt{2\pi\epsilon\Omega}} \exp\left(-\frac{\nu^2}{2\epsilon\Omega} \right).$$

From the definition:

$$P(q, t) = \left\langle \prod_\mu \delta(q^\mu(t) - q^\mu) \right\rangle$$

it follows that

$$P(q, t + \epsilon) = \left\langle \prod_\mu \delta\Big(q^\mu(t + \epsilon) - q^\mu \Big) \right\rangle.$$

Using Eq. 16, the following property of the Dirac delta function

$$\prod_\mu \delta(e^\mu_{\ i} A^{\ i}) = \frac{1}{\det(e^\mu_{\ i})} \prod_i \delta(A^i),$$

and the fact that $(\det(e^\mu_{\ i})) = 1/\sqrt{g}$ we obtain by averaging once more (see details in the Apprndix):

$$P(q, t + \epsilon) = \langle P(q, t + \epsilon) \rangle = \int d^D q \sqrt{g} H(\sigma, q) P(q, t)$$

where

$$H(\sigma, q) = \frac{1}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \exp(-\frac{1}{2\epsilon\Omega} g_{\mu\nu} Q^\mu Q^\nu). \tag{17}$$

Here

$$Q^\mu = \sigma^\mu - q^\mu + \frac{1}{2}\, e^\mu_{\ i} \nabla_\lambda e^\lambda_{\ i}. \qquad (18)$$

The final result reads

$$P(\vec{\sigma}, t+\epsilon) = \int \frac{d^D\sigma\sqrt{g}}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \exp(-\frac{1}{2\epsilon\Omega} g_{\mu\nu} Q^\mu Q^\nu) P(\vec{\sigma}, t) \qquad (19)$$

The meaning of this expression is clear. Take for example the case where $g_{\mu\nu} = Id$ i.e. a flat space. In this case the Laplace-Beltrami is simply the Laplacian and the well known solution which is a convolution with a Gaussian is reproduced here by our short time kernel. In the Perona-Malik equation [3] the spatial metric is locally proportional to the identity, where the proportionality depends on the absolute value of the gradient at the point. The short time kernel means in this case that there is a local decision on the radius of diffusion. Technically the variance of the Gaussian depends locally on $|\nabla P|$ and is equal for the two local coordinates. For the Beltrami flow with the induced metric the Gaussian is not symmetric and the diffusion is not symmetric in the two directions. We can think of it as if the diffusion is done according to ellipses whose direction , size and eccentricity is decided locally according to the partial derivatives. It is reasonable to expect, from this point of view, that the Beltrami flow performs a better adaptive smoothing processing.

The analysis presented in this section gives an approximation to the full short time kernel. The exact kernel for the 1D case is derived in [6]. The derivation of the 2D exact short time kernel is beyond the scope of this paper and will be presented elsewhere.

Yet another representation to the solution of the Fokker-Planck equation exists. The Feynman-Kac Path integral solution. The Probability distribution is given in this case as

$$P(\vec{q}'', t) = \int_{q(t')=q'}^{q(t'')=q''} \prod_t \big(dq(t)\sqrt{g(q(t))}\big) \exp(-S(q)) \qquad (20)$$

where

$$S(q) = \frac{1}{2}\int_{t'}^{t''} dt \big[\dot{q}_\mu g^{\mu\nu} \dot{q}_\nu - e^\lambda_\mu (\partial^\mu e^\rho_\lambda) \dot{q}_\rho\big]. \qquad (21)$$

A more general process that involves a drift term can be analyzed. The exact form can be find in [**?**]. These expression are not practical from a computational point of view but they are valuable in that they open the way to a refined analysis of correlations on the image by the method of Feynman diagrams.

5. summary and conclusions

We derive the Beltrami flow from a stochastic point of view. Given a Brownian motion (a Langevin equation) on the image manifold we can calculate the equation that governs the time evolution of the probability distribution of finding a "particle" that obeys the Langevin equation at a specific point on the manifold after time t. This equation is the Fokker-Planck equation associated with the given Brownian motion on the image manifold. Interpreting the gray-level intensity as proportional to the probability distribution we obtain the Beltrami flow.

The importance of this new point of view is in the ability in one hand to generalize in a systematic way by introducing drift term and on the other to derive a short time kernel to this highly non-linear equation. Note that the non-linearity enters the diffusion equation through the dependence of the metric on the intensity. The non-linear diffusion in its explicit form is therefore, strictly speaking, not derivable as a Fokker-Planck equation. The analysis presented in this paper gives an approximation to the full short time kernel. The exact kernel for the 1D case is derived in [6]. The derivation of the 2D exact short time kernel is beyond the scope of this paper and will be presented elsewhere.

The stochastic viewpoint enables us also to envisage new and exiting possibilities: Can we relate the stochastic nature of the of image formation process to the stochastic process which underlie the image manifold evolution? Can we have an efficient tool to study correlation between different parts of the image, or between different images of the same scene, using the path integral to compute Feynman diagrams?

These directions of research may lead to new understanding and new techniques in image processing and analysis.

Appendix

The noise $\nu^i(t)$ satisfies

$$\begin{aligned} \langle \nu_i(t) \rangle &= 0 \qquad (22) \\ \langle \nu_i(t)\nu_j(t) \rangle_\epsilon &= K_{ij}(t,t') = \delta_{ij}\eta_\epsilon(t-t') \qquad (23) \end{aligned}$$

where $\eta_\epsilon(t)$ is an even function which peaked around $t = 0$ and normalized according to $\int_{-\infty}^{\infty} \eta(t)dt = 1$. In the limit $\epsilon \to 0$ the function η_ϵ approaches the Dirac delta "function". the angle brackets denotes average with respect to the noise. The measure on the noise is given by

$$d\rho(\nu) = \frac{d\nu}{\sqrt{2\pi\Omega}} \exp\left(-\frac{1}{2\Omega}\int dt dt' \nu_i(t)(K^{-1})_{ij}(t,t')\nu_j(t')\right) ,$$

and the average for any function of ν is

$$\langle F(\nu) \rangle_\epsilon = \int d\rho(\nu) F(\nu).$$

and we denote by a bracket with no subscript the limit of $\epsilon \to 0$:

$$\langle F(\nu) \rangle = \lim_{\epsilon \to 0} \langle F(\nu) \rangle_\epsilon$$

Discretized Langevin equation: We discretize the time $t - t_0 = n\epsilon$ and the Langevin equation is:

$$q^\mu(t+\epsilon) = q^\mu(t) + \frac{1}{2}\, e^\mu_{\ i} \nabla_\nu e^\nu_{\ i}(q) + e^\mu_{\ i}(q)\nu_i\,.$$

The measure on the noise ν is

$$d\rho(\nu) = \frac{d\nu}{\sqrt{2\pi\epsilon\Omega}} \exp\left(-\frac{\nu^2}{2\epsilon\Omega}\right)\,.$$

From the definition:

$$P(q,t) = \left\langle \prod_\mu \delta(q^\mu(t) - q^\mu) \right\rangle$$

and therefore

$$\begin{aligned}
&P(q,t+\epsilon) = \left\langle \prod_\mu \delta\Big(q^\mu(t+\epsilon) - q^\mu\Big) \right\rangle \\
&= \int \frac{d^D\nu}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \exp(-\frac{\nu^2}{2\epsilon\Omega}) \\
&\times \prod_\mu \delta(q^\mu(t) + \frac{1}{2}\, e^\mu_{\ i}\nabla_\nu e^\nu_{\ i} + e^\mu_{\ i}\nu_i - q^\mu) \\
&= \int \frac{d^D\nu d^D\sigma}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \prod_\mu \delta(q^\mu(t) - \sigma^\mu) \int \frac{d^D\nu \exp(-\frac{\nu^2}{2\epsilon\Omega})}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \\
&\times \prod_\mu \delta\Big(e^\mu_{\ i}\Big(e^i_{\ \nu}(\sigma^\nu + \frac{1}{2}\, e^\nu_{\ j}\nabla_\lambda e^\lambda_{\ j} - q^\nu) + \nu_i\Big)\Big)
\end{aligned}$$

Averaging once more we obtain

$$\begin{aligned}
&P(q,t+\epsilon) = \langle P(q,t+\epsilon)\rangle \\
&= \int \frac{d^D\nu d^D\sigma}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} P(\sigma,t) \exp(-\frac{\nu^2}{2\epsilon\Omega}) \\
&\times \prod_\mu \delta\Big(e^\mu_{\ i}\Big(e^i_{\ \nu}(\sigma^\nu + \frac{1}{2}\, e^\nu_{\ j}\nabla_\lambda e^\lambda_{\ j} - q^\nu) + \nu_i\Big)\Big) \\
&= \int \frac{d^D\nu d^D\sigma}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \frac{P(\sigma,t)}{(\det(e^\mu_{\ i}))} \exp(-\frac{\nu^2}{2\epsilon\Omega}) \\
&\times \prod_i \delta\Big(e^i_{\ \nu}(\sigma^\nu + \frac{1}{2}\, e^\nu_{\ j}\nabla_\lambda e^\lambda_{\ j} - q^\nu) + \nu_i\Big) \\
&= \int d^D\sigma \sqrt{g} H(q,\sigma) P(\sigma,t)
\end{aligned}$$

where we used the following property of the Dirac delta function

$$\prod_\mu \delta(e^\mu_{\ i} A^{\ i}) = \frac{1}{\det(e^\mu_{\ i})} \prod_i \delta(A^i),$$

the fact that $(\det(e^\mu_{\ i})) = 1/\sqrt{g}$ and we define

$$H(\sigma,q) = \frac{1}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \exp(-\frac{1}{2\epsilon\Omega} g_{\mu\nu} d^\mu d^\nu) \tag{24}$$

where

$$Q^\mu = \sigma^\mu - q^\mu + \frac{1}{2}\, e^\mu_{\ i}\nabla_\lambda e^\lambda_{\ i}. \tag{25}$$

The final result reads

$$P(\vec{\sigma}, t+\epsilon) = \int \frac{d^D\sigma}{(2\pi\epsilon\Omega)^{\frac{D}{2}}} \sqrt{g} \exp(-\frac{1}{2\epsilon\Omega} g_{\mu\nu} Q^\mu Q^\nu) P(\vec{\sigma}, t) \tag{26}$$

References

[1] R. Durett and M. A. Pinski, "Geometry of Random Motion", Proceedings of the AMS-IMS-SIAM Joint Summer Research Conference 1987, *Contemporary Mathematics* Vol. 73 American Mathematical society.

[2] R. Kimmel R. Malladi and N. Sochen. Images as embedded maps and minimal surfaces: Movies, color, texture, and volumetric medical images. LBNL ,UC and TAUHEP Report LBNL-40490, UC-405, June 1997.

[3] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.* **12**, pp. 629-639, 1990.

[4] A. M. Polyakov. Quantum geometry of bosonic strings. *Physics Letters B*, 103B(3):207–210, 1981.

[5] N. Sochen, R. Kimmel, and R. Malladi. From high energy physics to low level vision. Report LBNL 39243, LBNL, UC Berkeley, CA 94720, August 1996.

[6] N. Sochen, R. Kimmel ans A. M. Bruckstein, Diffusions and Confusions in Signal and Image Processing Journal of Mathematical Imaging and Vision, in press

[7] N. Sochen R. Kimmel and R. Malladi. A general framework for low level vision. *IEEE Trans. IP*, 7:310, 1997.

[8] N Sochen and Y Y Zeevi, "Images as manifolds embedded in a spatial-feature non-Euclidean space", November 1998, EE-Technion report no. 1181.

[9] N. Sochen and Y.Y. Zeevi. Representation of Colored Images by Manifolds Embedded in Higher Dimensional Non-Euclidean Space *ICIP98*, 1998.

[10] J. Zinn-Justin, "Quantum field theory and critical phenomena", *Oxford University Press*, 1989.