Probability Density Under Transformation

Pramook Khungurn

December 11, 2019

1 Introduction

In creating an algorithm that samples points from some domain, a problem that always comes up is the following:

> Let A and B be sets,
> $p_A(\cdot)$ be a probability density on A, and
> f be a function from A to B.
> If one samples x from A according to p_A, then what is the probability density of $f(x)$?

This document discusses the solution to the above problem and its application to construction of sampling algorithm.

2 One-Dimensional Case

2.1 The Main Theorem

We first start with the simplest case where A and B are both subsets of the real line $\mathbb{R}$.

Let $x \in A$. The number $p_A(x)$ means that, in the infinitesimal interval $[x, x + \delta x)$, there exists $p_A(x)\delta x$ amount of "probability mass." Here, δx is a "differential quantity" such that $(\delta x)^2 = 0$.

Assume that f is continuous and infinitely differentiable. The function f sends the interval $[x, x + \delta x)$ to the interval $[f(x), f(x + \delta x))$. By Taylor expansion,

$$f(x + \delta x) = f(x) + f'(x)\delta x + O((\delta x)^2) = f(x) + f'(x)\delta x.$$

So, the resuling interval is $[f(x), f(x) + f'(x)\delta x)$, which as width $|f'(x)|\delta x$.

This means that the mass $p_A(x)\delta x$ gets distributed to an interval of width $f'(x)\delta x$. As a result:

$$\text{Density at point f(x)} = \frac{p_A(x)\delta x}{|f'(x)|\delta x} = \frac{p_A(x)}{|f'(x)|}.$$

This density is defined only when $f'(x) \neq 0$, which means that f is one-to-one in a neighborhood of x. As such, we have the following theorem.

Theorem 1. *Let A and B be subsets of $\mathbb{R}$, p_A be a probability density on A, $f : A \to B$ be continuous and differentiable and $f'(x) \neq 0$ for all $x \in A$. The induced probability density $p_B(\cdot)$ arisen from the process of sampling x according to p_A and then computing $f(x)$ is given by:*

$$p_B(f(x)) = \frac{p_A(x)}{|f'(x)|}.$$

2.2 The Inversion Method

The above theorem can be used to create sampling algorithm for any integrable density function on the real line from a uniformly random sample from the interval $[0, 1)$.

In this situation, $A = [0, 1)$ and $p_A(x) = 1$ for all $x \in A$. The density $p_B(\cdot)$ is given to us. We want to find function $f : A \to B$ such that, for any $x \in A$:

$$p_B(f(x)) = \frac{p_A(x)}{f'(x)} = \frac{1}{|f'(x)|}.$$

Multiply both sides by $f'(x)$, we have:

$$p_B(f(x))|f'(x)| = 1.$$

Let P_B be the CDF of p_B:

$$P_B(y) = \int_{-\infty}^{y} p_B(t)\ \mathrm{d}t.$$

We have that:

$$\{P_B(f(x))\}' = p_B(f(x))f'(x) = p_B(f(x))|f'(x)|$$

given that f is an increasing function. Let us assume that f is increasing for now. We have that

$$\{P_B(f(x))\}' = 1$$

Integrating both sides from $t = 0$ to $t = x$, we have:

$$\begin{aligned} \int_0^x \{P_B(f(t))\}'\ \mathrm{d}t &= \int_0^x 1\ \mathrm{d}t \\ P_B(f(x)) - P_B(f(0)) &= x. \end{aligned}$$

With the assumption that $f(0)$ should correspond to the lowest number in the set B, we can safely set $P_B(f(0)) = 0$. So,

$$\begin{aligned} P_B(f(x)) &= x \\ f(x) &= P_B^{-1}(x). \end{aligned}$$

The CDF is an increasing function, so is its inverse. Moreover, $P_B^{-1}(0)$ maps to the lowest number in the set B. So, it is a valid choice for f.

In other words, to generate a point on the real line with probability distribution p_B, simply apply the inverse of the CDF to a point x picked uniformly randomly from the interval $[0, 1)$.

2.3 Sampling from the Exponential Distribution

We present a simple application of the inversion method. The exponential distribution with parameter λ is defined on $[0, \infty)$ with

$$p(x) = \lambda e^{-\lambda x}.$$

The CDF is given by:

$$P(x) = 1 - e^{-\lambda x}.$$

So,

$$P^{-1}(y) = \ln(1 - y).$$

Hence, to sample x acoording to the exponential distribution, we simply set:

$$x := \ln(1 - \xi)$$

where ξ is a randomly and uniformly sampled from the interval $[0, 1)$.

3 Multi-Dimensional Case

3.1 The Main Theorem

Let $A, B \subseteq \mathbb{R}^n$, and $p_A(\cdot)$ be a probability density on A. Let $\mathbf{f}$ be given by:

$$\mathbf{f}(x_1, x_2, \dots, x_n) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) \end{bmatrix}$$

be a function from A to B. The induced probability distribution p_B arisen from the process of sampling a point $\mathbf{x} = (x_1, x_2, \dots, x_n)$ according to p_A can then computing $\mathbf{f}(\mathbf{x})$ can again be computed by finding the volume of the image of the interval

$$[x_1, x_1 + \delta x_1) \times [x_2, x_2 + \delta x_2) \times \cdots \times [x_n, x_n + \delta x_n).$$

This volume is given by:

$$|D\mathbf{f}(\mathbf{x})| \; \delta x_1 \delta x_2 \dots \delta x_n$$

where

$$D\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \cdots & \partial f_2/\partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1 & \partial f_n/\partial x_2 & \cdots & \partial f_n/\partial x_n \end{bmatrix}$$

where all the partial derivatives are evaluated at $\mathbf{x}$. Thus,

$$p_B(\mathbf{f}(\mathbf{x})) = \frac{p_A(\mathbf{x})}{|D\mathbf{f}(\mathbf{x})|}.$$

Notice that $|D\mathbf{f}(\mathbf{x})|$ is the factor that shows up when we perform change of variables during an integration.

In two-dimensional space, we may write:

$$\mathbf{f}(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \end{bmatrix}.$$

In this case:

$$|D\mathbf{f}(u, v)| = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix}.$$

Thus,

$$p_B(x, y) = \frac{p_A(u, v)}{\begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix}}.$$

3.2 The Polar Coordinate Transform

The polar coordinate transforms two numbers (r, ϕ) to a point (x, y) on the plane as follows:

$$x = r \cos \phi$$
$$y = r \sin \phi,$$

which gives:

$$\frac{\partial x}{\partial r} = \cos\phi$$
$$\frac{\partial x}{\partial \phi} = -r\sin\phi$$
$$\frac{\partial y}{\partial r} = \sin\theta$$
$$\frac{\partial y}{\partial \phi} = r\cos\phi$$

So, if we sample a polar coordinate (r, ϕ) with probability distribution p_A, then the distribution p_B of the point (x, y) is given by:

$$p_B(x,y) = \frac{p_A(r,\phi)}{\begin{vmatrix}\cos\phi & -r\sin\phi \\ \sin\theta & r\cos\phi\end{vmatrix}} = \frac{p_A(r,\phi)}{r\cos^2\phi + r\sin^2\phi} = \frac{p_A(r,\phi)}{r}.$$

3.3 Sampling Uniformly from the Unit Disk

The unit disk is given by the polar coordinates in the set $[0, 1] \times [0, 2\pi)$. How should we be sampling the polar coordinates so that the resulting point distribution is uniform on the disk?

In our case, we have that $p_B(x, y) = 1/\pi$. So, we want p_A such that:

$$\frac{1}{\pi} = \frac{p_A(r,\phi)}{r}$$
$$p_A(r,\phi) = \frac{r}{\pi}.$$

A common strategy is to sample r and ϕ independently so that $p_A(r, \phi) = p_r(r)p_\phi(\phi)$. Moreover, we shall sample ϕ uniformly from the interval $[0, 2\pi)$ so that $p_\phi(\phi) = 1/(2\pi)$. Thus,

$$p_r(r) = 2r.$$

The above distribution can be sampled with the inversion method. The CDF is given by:

$$P_r(r) = \int_0^r 2r' \, \mathrm{d}r' = [r'^2]_0^r = r^2.$$

The inverse CDF is then:

$$P_r^{-1}(t) = \sqrt{t}.$$

So, we can sample points uniformly from the unit disk by setting:

$$r := \sqrt{\xi_1}$$
$$\phi = 2\pi\xi_2$$

where ξ_1 and ξ_2 are two independent random samples chosen uniformly from the interval $[0, 1)$.

3.4 Sampling Uniformly from a Triangle

Suppose we have a triangle in a plane with point $A = (x_A, y_A)$, $B = (x_B, y_B)$, $C = (x_C, y_C)$. Let us assume further that $(B - A) \times (C - A)$ is pointing in the positive z-direction so that:

$$\text{area}(ABC) = \frac{1}{2}\|(B-A)\times(C-A)\| = \frac{1}{2}\begin{vmatrix} x_B - x_A & x_C - x_A \\ y_B - y_A & y_C - y_A. \end{vmatrix}$$

We wish to find a transformation $\mathbf{f}$ that takes a point (u, v) uniformly and randomly picked from the rectangle $[0, 1)^2$ so that the distribution of $(x, y) = \mathbf{f}(u, v)$ is uniform on the triangle ABC. In this setting, we have that $p_A(u, v) = 1$, and $p_B(x, y) = 1/\text{area}(ABC)$. In other words,

$$\begin{aligned}
\frac{1}{\text{area}(ABC)} &= \frac{1}{|D\mathbf{f}(u, v)|} \\
|D\mathbf{f}(u, v)| &= \frac{1}{2} \begin{vmatrix} x_B - x_A & x_C - x_A \\ y_B - y_A & y_C - y_A \end{vmatrix}.
\end{aligned}$$

One way to generate a point on a triangle is to generate barycentric coordinates (α, β, γ) such that $0 \leq \alpha, \beta, \gamma \leq 1$ and $\alpha + \beta + \gamma = 1$. Then, we can get a point on the triangle by computing

$$\begin{aligned}
(x, y) &= \alpha A + \beta B + \gamma C \\
&= (1 - \beta - \gamma)A + \beta B + \gamma C \\
&= A + (B - A)\beta + (C - A)\gamma.
\end{aligned}$$

In other words,

$$\begin{aligned}
x &= x_A + (x_B - x_A)\beta + (x_C - x_A)\gamma \\
y &= y_A + (y_B - y_A)\beta + (y_C - y_A)\gamma.
\end{aligned}$$

Our task is to figure out what β and γ are as functions of u and v.

We have that

$$\begin{aligned}
\frac{\partial x}{\partial u} &= (x_B - x_A)\frac{\partial \beta}{\partial u} + (x_C - x_A)\frac{\partial \gamma}{\partial u} \\
\frac{\partial x}{\partial v} &= (x_B - x_A)\frac{\partial \beta}{\partial v} + (x_C - x_A)\frac{\partial \gamma}{\partial v} \\
\frac{\partial y}{\partial u} &= (y_B - y_A)\frac{\partial \beta}{\partial u} + (y_C - y_A)\frac{\partial \gamma}{\partial u} \\
\frac{\partial y}{\partial v} &= (y_B - y_A)\frac{\partial \beta}{\partial v} + (y_C - y_A)\frac{\partial \gamma}{\partial v}.
\end{aligned}$$

So, the matrix $D\mathbf{f}(u, v)$ is given by:

$$\begin{aligned}
D\mathbf{f}(u, v) &= \begin{bmatrix} (x_B - x_A)\frac{\partial \beta}{\partial u} + (x_C - x_A)\frac{\partial \gamma}{\partial u} & (x_B - x_A)\frac{\partial \beta}{\partial v} + (x_C - x_A)\frac{\partial \gamma}{\partial v} \\ (y_B - y_A)\frac{\partial \beta}{\partial u} + (y_C - y_A)\frac{\partial \gamma}{\partial u} & (y_B - y_A)\frac{\partial \beta}{\partial v} + (y_C - y_A)\frac{\partial \gamma}{\partial v} \end{bmatrix} \\
&= \begin{bmatrix} x_B - x_A & x_C - x_A \\ y_B - y_A & y_C - y_A \end{bmatrix} \begin{bmatrix} \partial\beta/\partial u & \partial\beta/\partial v \\ \partial\gamma/\partial u & \partial\gamma/\partial v \end{bmatrix}.
\end{aligned}$$

Thus,

$$\begin{aligned}
|D\mathbf{f}(u, v)| &= \begin{vmatrix} x_B - x_A & x_C - x_A \\ y_B - y_A & y_C - y_A \end{vmatrix} \begin{vmatrix} \partial\beta/\partial u & \partial\beta/\partial v \\ \partial\gamma/\partial u & \partial\gamma/\partial v \end{vmatrix} \\
\frac{1}{2} \begin{vmatrix} x_B - x_A & x_C - x_A \\ y_B - y_A & y_C - y_A \end{vmatrix} &= \begin{vmatrix} x_B - x_A & x_C - x_A \\ y_B - y_A & y_C - y_A \end{vmatrix} \begin{vmatrix} \partial\beta/\partial u & \partial\beta/\partial v \\ \partial\gamma/\partial u & \partial\gamma/\partial v \end{vmatrix} \\
\frac{1}{2} &= \begin{vmatrix} \partial\beta/\partial u & \partial\beta/\partial v \\ \partial\gamma/\partial u & \partial\gamma/\partial v \end{vmatrix} \\
\frac{\partial \beta}{\partial u}\frac{\partial \gamma}{\partial v} - \frac{\partial \beta}{\partial v}\frac{\partial \gamma}{\partial u} &= \frac{1}{2}.
\end{aligned}$$

What should β and γ be as functions of u and v? We have the constraint that $0 \leq \beta + \gamma \leq 1$. This condition is satisfied if we let

$$\begin{aligned}
\beta &= g(u)(1 - v) \\
\gamma &= g(u)v
\end{aligned}$$

where $g(u)$ is a function such that $0 \leq g(u) \leq 1$. With this choice of β and γ, we have that

$$\frac{1}{2} = \frac{\partial\beta}{\partial u}\frac{\partial\gamma}{\partial v} - \frac{\partial\beta}{\partial v}\frac{\partial\gamma}{\partial u} = [g'(u)(1-v)]g(u) - [-g(u)][g'(u)v] = g(u)g'(u).$$

It remains to find the function g with makes the above equation holds:

$$\begin{aligned} g\frac{\mathrm{d}g}{\mathrm{d}u} &= \frac{1}{2} \\ 2g\ \mathrm{d}g &= \mathrm{d}u \\ \int 2g\ \mathrm{d}g &= \int \mathrm{d}u \\ g^2 &= u \\ g &= \sqrt{u}. \end{aligned}$$

Hence, a uniform distribution of points on triangle ABC can be generated by computing:

$$(1 - \sqrt{u}(1-v) - \sqrt{u}v)A + \sqrt{u}(1-v)B + \sqrt{u}vC$$

where (u,v) is randomly and uniformly sampled from the rectangle $[0,1)^2$.

4 Dealing with 3D Manifolds

4.1 The Main Theorem

Suppose that we have a differentiable function $\mathbf{f}$ that maps a set $A \subseteq \mathbb{R}^2$ to a surface $B \subseteq \mathbb{R}^3$. We shall write:

$$\mathbf{f}(u,v) = \begin{bmatrix} f_x(u,v) \\ f_y(u,v) \\ f_z(u,v) \end{bmatrix}.$$

Again, let p_A be a probability distribution on A. Given point $(u,v) \in A$, consider the rectangle $[u + \delta u) \times [v + \delta v)$, which has area $\delta u \delta v$. This rectangle has probability mass $p_A(u,v)\delta u \delta v$ in it.

We have that:

$$\begin{aligned} (u,v) &\mapsto \mathbf{f}(u,v) \\ (u+\delta u, v) &\mapsto \mathbf{f}(u+\delta u, v) = \mathbf{f}(u,v) + \mathbf{f}_u(u,v)\delta u \\ (u, v+\delta v) &\mapsto \mathbf{f}(u, v+\delta v) = \mathbf{f}(u,v) + \mathbf{f}_v(u,v)\delta v \\ (u+\delta u, v+\delta v) &\mapsto \mathbf{f}(u+\delta u, v+\delta v) = \mathbf{f}(u,v) + \mathbf{f}_u(u,v)\delta u + \mathbf{f}_v(u,v)\delta v \end{aligned}$$

where

$$\mathbf{f}_u(u,v) = \begin{bmatrix} \frac{\partial f_x}{\partial u}(u,v) \\ \frac{\partial f_y}{\partial u}(u,v) \\ \frac{\partial f_z}{\partial u}(u,v) \end{bmatrix}, \text{ and}$$

$$\mathbf{f}_v(u,v) = \begin{bmatrix} \frac{\partial f_x}{\partial v}(u,v) \\ \frac{\partial f_y}{\partial v}(u,v) \\ \frac{\partial f_z}{\partial v}(u,v) \end{bmatrix}.$$

In other words, the rectangle gets mapped to a parallelogram with sides defined by the vector $\mathbf{f}_u(u,v)\delta u$ and $\mathbf{f}_v(u,v)\delta v$. The area of this parallelogram is given by:

$$\|\mathbf{f}_u(u,v)\delta u \times \mathbf{f}_u(u,v)\delta v\| = \|\mathbf{f}_u(u,v) \times \mathbf{f}_u(u,v)\|\delta u \delta v$$

(Since the notation is getting a little unwieldy, let us drop the (u, v) arguments from the function from now on.) To compute the cross product, we make use of the following identity:

$$\|\mathbf{a} \times \mathbf{b}\|^2 + (\mathbf{a} \cdot \mathbf{b})^2 = (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b}).$$

So,

$$\begin{aligned} \|\mathbf{f}_u \times \mathbf{f}_v\|^2 &= (\mathbf{f}_u \cdot \mathbf{f}_u)(\mathbf{f}_v \cdot \mathbf{f}_v) - (\mathbf{f}_u \cdot \mathbf{f}_v)^2 \\ \|\mathbf{f}_u \times \mathbf{f}_v\| &= \sqrt{(\mathbf{f}_u \cdot \mathbf{f}_u)(\mathbf{f}_v \cdot \mathbf{f}_v) - (\mathbf{f}_u \cdot \mathbf{f}_v)^2}. \end{aligned}$$

Define

$$\begin{aligned} E(u, v) &= \mathbf{f}_u(u, v) \cdot \mathbf{f}_u(u, v) \\ F(u, v) &= \mathbf{f}_u(u, v) \cdot \mathbf{f}_v(u, v) \\ G(u, v) &= \mathbf{f}_v(u, v) \cdot \mathbf{f}_v(u, v). \end{aligned}$$

We have that:

$$\text{area of parallelogram} = \|\mathbf{f}_u \times \mathbf{f}_v\| = \sqrt{EG - F^2}.$$

In differential geometry, E, F, and G are called the *coefficients of the first fundamental form*.

As a result, we have that the induced probability distribution is given by:

$$p_B(\mathbf{f}(u, v)) = \frac{p_A(u, v)\delta u \delta v}{\|\mathbf{f}_u \times \mathbf{f}_v\| \delta u \delta v} = \frac{p_A(u, v)}{\sqrt{EG - F^2}}.$$

4.2 The Spherical Coordinate Transform

The spherical coordinate is the transformation from $(\theta, \phi) \in (0, \pi) \times [0, 2\pi)$ to a point ω on a 3D sphere S^2 with:

$$\omega = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}.$$

We then have that:

$$\begin{aligned} \omega_\theta &= \begin{bmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{bmatrix}, \\ \omega_\phi &= \begin{bmatrix} -\sin\theta\sin\phi \\ \sin\theta\cos\phi \\ 0 \end{bmatrix}. \end{aligned}$$

So,

$$\begin{aligned} E &= \cos^2\theta\cos^2\phi + \cos^2\theta\sin^2\phi + \sin^2\theta \\ &= \cos^2\theta + \sin^2\theta \\ &= 1 \\ F &= -\cos\theta\cos\phi\sin\theta\sin\phi + \cos\theta\sin\phi\sin\theta\cos\phi + 0 \\ &= 0 \\ G &= \sin^2\theta\sin^2\phi + \sin^2\theta\cos^2\phi \\ &= \sin^2\theta \\ \sqrt{EG - F^2} &= \sqrt{\sin^2\theta} = |\sin\theta|. \end{aligned}$$

The inducted probability distribution is given by:

$$p_B(\omega(\theta,\phi)) = \frac{p_A(\theta,\phi)}{|\sin\theta|}.$$

However, since $\theta \in (0,\theta)$, we have that $\sin\theta > 0$. So, we can write:

$$p_B(\omega(\theta,\phi)) = \frac{p_A(\theta,\phi)}{\sin\theta}.$$

4.3 Uniformly Sampling a Sphere

We will use the identity to construct a sampling algorithm to sample a point on the unit sphere uniformly. The idea is to pick a probability distribution p_A on $(\theta,\phi) \in (0,\pi)\times[0,2\pi)$ such that the induced probability distribution p_B is the constant distribution $1/(4\pi)$. In other words:

$$\frac{1}{4\pi} = \frac{p_A(\theta,\phi)}{\sin\theta}.$$

In other words:

$$p_A(\theta,\phi) = \frac{\sin\theta}{4\pi}.$$

A common strategy is to sample ϕ independenty from θ so that $p_A(\theta,\phi) = p_\theta(\theta)p_\phi(\phi)$. Moreover, let us sample ϕ uniformly from $[0,2\pi)$ so that $p_\phi(\phi) = 1/(2\pi)$. In other words,

$$\begin{aligned}\frac{p_\theta(\theta)}{2\pi} &= \frac{\sin\theta}{4\pi}\\ p_\theta(\theta) &= \frac{\sin\theta}{2}.\end{aligned}$$

We can sample $p_\theta(\theta)$ using the inversion method. The CDF of p_Θ is given by:

$$P_\theta(\theta) = \frac{1}{2}\int_0^\theta \sin\theta'\ \mathrm{d}\theta' = \frac{1}{2}[-\cos\theta']_0^\theta = \frac{\cos 0 - \cos\theta}{2} = \frac{1-\cos\theta}{2}.$$

So, the inverse function is given by:

$$P_\theta^{-1}(u) = \cos^{-1}(1-2u).$$

In conclusion, we compute θ and ϕ as:

$$\begin{aligned}\theta &:= \cos^{-1}(1-2\xi_0)\\ \phi &:= 2\pi\xi_1\end{aligned}$$

where ξ_0, ξ_1 are two independent random numbers sampled uniformly from the interval $[0,1)$.

Notice, however, that if the end goal is to get a point ω, there is no need to compute θ because θ never appears directly in the expression for ω. More specifically,

$$\omega = \begin{bmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{bmatrix} = \begin{bmatrix}\sqrt{1-(1-2\xi_0)^2}\cos\phi\\ \sqrt{1-(1-2\xi_0)^2}\sin\phi\\ 1-2\xi_0\end{bmatrix}.$$

4.4 Sampling a Cosine-Weighted Hemisphere

In this section, we want to sample the z-positive unit hemisphere such that the probability density being proportional to $\cos\theta$ at each point. In this case:

$$\frac{\cos\theta}{C} = \frac{p_A(\theta,\phi)}{\sin\theta}$$
$$\frac{1}{C}\cos\theta\sin\theta = p_A(\theta,\phi),$$

where C is the constant such that $\frac{\cos\theta}{C}$ is a probability distribution on the sphere.

Again, we sample θ and ϕ independently with ϕ being uniform in $[0, 2\pi)$. So,

$$\frac{2\pi}{C}\cos\theta\sin\theta = p_\theta(\theta).$$

The CDF then is given by:

$$P_\theta(\theta) = \frac{2\pi}{C}\int_0^\theta \cos\theta'\sin\theta'\ \mathrm{d}\theta' = \frac{2\pi}{C}\left[-\frac{\cos^2\theta'}{2}\right]_0^\theta = \frac{\pi}{C}(1-\cos^2\theta).$$

To determine C, note that $P_\theta(\pi/2) = 1$, so

$$1 = \frac{\pi}{C}(1-\cos^2\frac{\pi}{2}) = \frac{\pi}{C}.$$

In other words, $C = \pi$, and $P_\theta(\theta) = 1 - \cos^2\theta$.

Hence, we can sample the cosine-weighted hemisphere by setting:

$$\cos\theta := \sqrt{1-\xi_0}$$
$$\phi := 2\pi\xi_1.$$

4.5 From Area to Solid Angle

When shading from an area light source, a way to sample the incoming light direction is to sample a point on the light source's surface with some probability density p_A and then convert the vector from the shaded point to the sampled point to a unit vector ω. In this section, we find the relation between p_A and the induced probability density.

For simplicity, let us say that the shaded point is at the origin and lying on the xy-plane so that the normal is the z-axis. Let $\mathbf{r} = (r_x, r_y, r_z)$ denote the point on the light source. Let $\mathbf{n}$ be the normal at $\mathbf{r}$, and let $\mathbf{s}$ and $\mathbf{t}$ be the basis of the tangent plane at $\mathbf{r}$ in such a way that $(\mathbf{s}, \mathbf{t}, \mathbf{n})$ is an orthonormal basis. The tangent plane is the set

$$\{\mathbf{r} + u\mathbf{s} + v\mathbf{t} \mid (u,v) \in \mathbb{R}^2\}.$$

The function $\mathbf{f}$ that maps the tangent plane to the direction is given by:

$$\omega = \mathbf{f}(u,v) = \frac{\mathbf{r}+u\mathbf{s}+v\mathbf{t}}{\|\mathbf{r}+u\mathbf{s}+v\mathbf{t}\|}$$

Hence, using Lemma 2 (proven in the appendix), we have:

$$\mathbf{f}_u(u,v) = \frac{\mathbf{s}}{\|\mathbf{r}+u\mathbf{s}+v\mathbf{t}\|} - \frac{\mathbf{r}+u\mathbf{s}+v\mathbf{t}}{\|\mathbf{r}+u\mathbf{s}+v\mathbf{t}\|^3}(\mathbf{r}\cdot\mathbf{s}+u)$$
$$\mathbf{f}_v(u,v) = \frac{\mathbf{t}}{\|\mathbf{r}+u\mathbf{s}+v\mathbf{t}\|} - \frac{\mathbf{r}+u\mathbf{s}+v\mathbf{t}}{\|\mathbf{r}+u\mathbf{s}+v\mathbf{t}\|^3}(\mathbf{r}\cdot\mathbf{t}+v)$$

At $(u, v) = (0, 0)$, we have that

$$\mathbf{f}_u(0,0) = \frac{\mathbf{s}}{\|\mathbf{r}\|} - \frac{\mathbf{r}}{\|\mathbf{r}\|^3}(\mathbf{r}\cdot\mathbf{s}) = \frac{\mathbf{s}\|\mathbf{r}\|^2 - \mathbf{r}(\mathbf{r}\cdot\mathbf{s})}{\|\mathbf{r}\|^3}$$
$$\mathbf{f}_v(0,0) = \frac{\mathbf{t}}{\|\mathbf{r}\|} - \frac{\mathbf{r}}{\|\mathbf{r}\|^3}(\mathbf{r}\cdot\mathbf{t}) = \frac{\mathbf{t}\|\mathbf{r}\|^2 - \mathbf{r}(\mathbf{r}\cdot\mathbf{t})}{\|\mathbf{r}\|^3}$$

So,

$$E = \frac{\|\mathbf{r}\|^4 - 2\|\mathbf{r}^2\|(\mathbf{r}\cdot\mathbf{s})^2 + \|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{s})}{\|\mathbf{r}\|^6} = \frac{\|\mathbf{r}\|^4 - \|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{s})^2}{\|\mathbf{r}\|^6} = \frac{\|\mathbf{r}\|^2 - (\mathbf{r}\cdot\mathbf{s})^2}{\|\mathbf{r}\|^4}$$
$$F = -\frac{\|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{s})(\mathbf{r}\cdot\mathbf{t})}{\|\mathbf{r}\|^6} = -\frac{(\mathbf{r}\cdot\mathbf{s})(\mathbf{r}\cdot\mathbf{t})}{\|\mathbf{r}\|^4}$$
$$G = \frac{\|\mathbf{r}\|^2 - (\mathbf{r}\cdot\mathbf{t})^2}{\|\mathbf{r}\|^4}$$

Next,

$$\begin{aligned} EG - F^2 &= \frac{\|\mathbf{r}\|^4 - \|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{s})^2 - \|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{t})^2 + (\mathbf{r}\cdot\mathbf{s})^2(\mathbf{r}\cdot\mathbf{t})^2}{\|\mathbf{r}^8\|} - \frac{(\mathbf{r}\cdot\mathbf{s})^2(\mathbf{r}\cdot\mathbf{t})^2}{\|\mathbf{r}^8\|} \\ &= \frac{\|\mathbf{r}\|^4 - \|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{s})^2 - \|\mathbf{r}\|^2(\mathbf{r}\cdot\mathbf{t})^2}{\|\mathbf{r}^8\|} \\ &= \frac{1}{\|\mathbf{r}\|^4}\left[1 - \left(\frac{\mathbf{r}}{\|\mathbf{r}\|}\cdot\mathbf{s}\right)^2 - \left(\frac{\mathbf{r}}{\|\mathbf{r}\|}\cdot\mathbf{t}\right)^2\right] \\ &= \frac{1}{\|\mathbf{r}\|^4}[1 - (\hat{\mathbf{r}}\cdot\mathbf{s})^2 - (\hat{\mathbf{r}}\cdot\mathbf{t})^2] \end{aligned}$$

where $\hat{\mathbf{r}}$ is the unit vector in the direction of $\mathbf{r}$. Because $\mathbf{s}$, $\mathbf{t}$, $\mathbf{n}$ forms an orthonormal basis and $\|\hat{\mathbf{r}}\| = 1$, we have that

$$1 = \|\hat{\mathbf{r}}\|^2 = (\hat{\mathbf{r}}\cdot\mathbf{s})^2 + (\hat{\mathbf{r}}\cdot\mathbf{t})^2 + (\hat{\mathbf{r}}\cdot\mathbf{n})^2.$$

So,

$$EG - F^2 = \frac{1}{\|\mathbf{r}\|^4}[1 - (\hat{\mathbf{r}}\cdot\mathbf{s})^2 - (\hat{\mathbf{r}}\cdot\mathbf{t})^2] = \frac{1}{\|\mathbf{r}\|^4}(\hat{\mathbf{r}}\cdot\mathbf{n})^2$$

Thus,

$$\sqrt{EG - F^2} = \sqrt{\frac{(\hat{\mathbf{r}}\cdot\mathbf{n})^2}{\|\mathbf{r}\|^4}} = \frac{|\hat{\mathbf{r}}\cdot\mathbf{n}|}{\|\mathbf{r}\|^2}.$$

In conclusion,

$$p_B(\mathbf{f}(\mathbf{r})) = \frac{\|\mathbf{r}^2\|}{|\hat{\mathbf{r}}\cdot\mathbf{n}|}p_A(\mathbf{r}) = \frac{\|\mathbf{r}^2\|}{|\cos\theta|}p_A(\mathbf{r})$$

4.6 Area Under Linear Transformation

Suppose we sample a point $\mathbf{r}$ on a surface with probability distribution $p_A(\mathbf{r})$. Then, we transform it to $\mathbf{q} = M\mathbf{r}$ where M is a linear transformation. What is the probability density $p_B(\mathbf{q})$?

At $\mathbf{r}$, let the tangent space be spanned by unit tangent vectors $\mathbf{s}$ and $\mathbf{t}$. The normal vector $\mathbf{n} = \mathbf{s} \times \mathbf{t}$, together with the two tangent vectors, forms an orthonomal coordinate system. The tangent plane is thus given by:

$$\tilde{\mathbf{r}}(u, v) = \mathbf{r} + u\mathbf{s} + v\mathbf{t}.$$

The image of the tangent plane under the transformation is given by:

$$\tilde{\mathbf{q}}(u, v) = M\tilde{\mathbf{r}}(u, v) = M\mathbf{r} + uM\mathbf{s} + vM\mathbf{t}.$$

So,

$$\begin{aligned} \tilde{\mathbf{q}}_u &= M\mathbf{s} \\ \tilde{\mathbf{q}}_v &= M\mathbf{t}. \end{aligned}$$

The area of the parallelogram after transformation is ths:

$$\|M\mathbf{s} \times M\mathbf{t}\|.$$

Using the identity

$$M\mathbf{a} \times M\mathbf{b} = \det(M)M^{-T}(\mathbf{a} \times \mathbf{b}),$$

we have that:

$$\|M\mathbf{s} \times M\mathbf{t}\| = |\det(M)| \|M^{-T}(\mathbf{s} \times \mathbf{t})\| = |\det(M)| \|M^{-T}\mathbf{n}\|.$$

As a result,

$$p_B(M\mathbf{r}) = \frac{p_A(\mathbf{r})}{|\det(M)| \|M^{-T}\mathbf{n}\|}.$$

4.7 Solid Angle Under Linear Transformation

Consider the following process. We sample a direction ω_A acoording to probability density $p_A(\omega_A)$. Then, we apply a linear transformation M and normalize it to get a direction

$$\omega_B = \frac{M\omega_A}{\|M\omega_A\|}.$$

What is the probability density $p_B(\omega_B)$?

The probability can be computed by considering the whole transformation in 2 steps: linear transformation and then normalizing. Namely,

$$\begin{aligned} \mathbf{r}_C &= M\omega_A \\ \omega_B &= \frac{\mathbf{r}_C}{\|\mathbf{r}_C\|}. \end{aligned}$$

In the first step, we sample a point from the unit sphere, and the unit sphere gets transformed into some ellipsoid. From Section 4.6, we have that:

$$p_C(\mathbf{r}_C) = \frac{p_A(\omega_A)}{|\det(M)| \|M^{-T}\omega_A\|}.$$

Note that we use ω_A as $\mathbf{n}$ because the normal vector of the unit sphere at any point is that point itself.

Now, we get a point $\mathbf{r}_C$ on an ellipsoid. The normal vector at $\mathbf{n}_C$ at $\mathbf{r}_C$ is:

$$\mathbf{n}_C = \frac{M^{-T}\omega_A}{\|M^{-T}\omega_A\|}.$$

The unit vector from the origin to $\mathbf{r}_C$ is given by

$$\frac{\mathbf{r}_C}{\|\mathbf{r}_C\|} = \frac{M\omega_A}{\|M\omega_A\|}.$$

Using 4.5, we have:

$$\begin{aligned}
p_B(\omega_B) &= \frac{\|\mathbf{r}_C\|^2}{\cos\theta} p_C(\mathbf{r}_C) \\
&= \frac{\|M\omega_A\|^2}{|\frac{M\omega_A}{\|M\omega_A\|} \cdot \frac{M^{-T}\omega_A}{\|M^{-T}\omega_A\|}|} \frac{1}{|\det(M)|\|M^{-T}\omega_A\|} p_A(\omega_A) \\
&= \frac{\|M\omega_A\|^3}{|\omega_A^T M^{-1} M\omega_A|} \frac{1}{|\det(M)|} p_A(\omega_A) \\
&= \frac{\|M\omega_A\|^3}{|\det(M)|\|\omega_A\|^2} p_A(\omega_A) \\
&= \frac{\|M\omega_A\|^3}{|\det(M)|} p_A(\omega_A).
\end{aligned}$$

4.8 The Hair Coordinate System Transform

The hair coordinate system maps $(\theta, \phi) \in (\pi/2, \pi/2) \times [0, 2\pi)$ to a sphere with the following transformation function:

$$\omega = \begin{bmatrix} \sin\theta \\ \cos\theta\cos\phi \\ \cos\theta\sin\phi \end{bmatrix}.$$

So,

$$\begin{aligned}
\omega_\theta &= \begin{bmatrix} \cos\theta \\ -\sin\theta\cos\phi \\ -\sin\theta\sin\phi \end{bmatrix} \\
\omega_\phi &= \begin{bmatrix} 0 \\ -\cos\theta\sin\phi \\ \cos\theta\cos\phi \end{bmatrix} \\
E &= \cos^2\theta + \sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi = 1 \\
F &= \sin\theta\cos\theta\cos\phi\sin\phi - \sin\theta\cos\theta\cos\phi\sin\phi = 0 \\
G &= \cos^2\theta\sin^2\phi + \cos^2\theta\cos^2\phi = \cos^2\theta \\
\sqrt{EG - F^2} &= \sqrt{\cos^2\theta - 0} = |\cos\theta|.
\end{aligned}$$

However, since $\theta \in (-\pi/2, \pi/2)$, we have that $\cos\theta > 0$, so

$$\sqrt{EG - F^2} = \cos\theta.$$

So, the probability density transformation formula is:

$$p_B(\omega(\theta, \phi)) = \frac{p_A(\theta, \phi)}{\cos\theta}.$$

4.9 Sampling for Diffuse Hair

In this section, we want to sample the sphere so that $p_B(\omega) \propto \cos\theta$. Applying the main theorem in this section, we have:

$$\begin{aligned} \frac{\cos\theta}{C} &= \frac{p_A(\theta,\phi)}{\cos\theta} \\ p_A(\theta,\phi) &= \frac{\cos^2\theta}{C}. \end{aligned}$$

Again, we sample ϕ uniformly from $[0, 2\pi)$, and then sample θ independently from ϕ. So,

$$\begin{aligned} p_\theta(\theta) &= \frac{2\pi}{C}\cos^2\theta \\ P_\theta(\theta) &= \frac{2\pi}{C}\int_{-\pi/2}^{\theta} \cos^2\theta' \ \mathrm{d}\theta' \\ &= \frac{2\pi}{C}\left[\frac{\theta' + \sin\theta'\cos\theta'}{2}\right]_{-\pi/2}^{\theta} \\ &= \frac{\pi}{C}\left[\theta' + \frac{\sin(2\theta')}{2}\right]_{-\pi/2}^{\theta} \\ &= \frac{\pi}{C}\left(\theta + \frac{\sin(2\theta)}{2} + \frac{\pi}{2}\right). \end{aligned}$$

To find C, we note that $P_\theta(\pi/2) = 1$, so

$$1 = \frac{\pi}{C}\left(\frac{\pi}{2} + 0 + \frac{\pi}{2}\right) = \frac{\pi^2}{C}$$

So, $C = \pi^2$, and

$$P_\theta(\theta) = \frac{1}{\pi}\left(\theta + \frac{\sin(2\theta)}{2} + \frac{\pi}{2}\right).$$

The above function cannot be inverted symbolically. So, in Mitsuba's implementation, they solve for it using Brent's method.

5 Radiance Under Linear Transformation

Suppose you have an area light source. Let $\mathbf{x}$ be a point on that surface, and let $L_e(\mathbf{x},\omega)$ be the outgoing radiance distribution at $\mathbf{x}$. If the whole light source is transformed by a linear transformation M, what is the outgoing radiance from the same point in the same direction after transformation calculated in such a way that energy is conserved?

For example, if we have an area light source that emits a constant radiance, say L, in all direction. If this light source is scaled uniformly by a factor of 2 in all directions, then the surface area of the scaled light source is 4 times that of the original one. As such, to conserve energy, the radiance from each point has to be $L/4$. We would like to find an expression for the radiance after transformation that correctly takes into account this behavior.

Let $\mathbf{n}$ be the normal at $\mathbf{x}$. Then, the power of the radiance along the ray from $\mathbf{x}$ in direction ω is given by:

$$L(\mathbf{x},\omega)(\omega\cdot\mathbf{n}) \ \mathrm{d}A(\mathbf{x})\mathrm{d}\sigma(\omega)$$

where A is the area measure, and σ is the solid angle measure.

After transformation, we get:

$$\begin{aligned}\tilde{\mathbf{x}} &= M\mathbf{x}\\ \tilde{\omega} &= \frac{M\omega}{\|M\omega\|}\\ \tilde{\mathbf{n}} &= \frac{M^{-T}\mathbf{n}}{\|M^{-T}\mathbf{n}\|}.\end{aligned}$$

The power of the radiance along the ray from $\tilde{\mathbf{x}}$ in direction $\tilde{\omega}$ is:

$$L(\tilde{\mathbf{x}}, \tilde{\omega})(\tilde{\omega} \cdot \tilde{\mathbf{n}})\ \mathrm{d}A(\tilde{\mathbf{x}})\mathrm{d}\sigma(\tilde{\omega}).$$

Because we want to conserve energy, we have that:

$$L(\tilde{\mathbf{x}}, \tilde{\omega})(\tilde{\omega} \cdot \tilde{\mathbf{n}})\ \mathrm{d}A(\tilde{\mathbf{x}})\mathrm{d}\sigma(\tilde{\omega}) = L(\mathbf{x}, \omega)(\omega \cdot \mathbf{n})\ \mathrm{d}A(\mathbf{x})\mathrm{d}\sigma(\omega).$$

Using the knowledge from Section 4.6,

$$\mathrm{d}A(\tilde{\mathbf{x}}) = |\det(M)|\|M^{-T}\mathbf{n}\|\ \mathrm{d}A(\mathbf{x}).$$

Moreover, from Section 4.7, we have:

$$\mathrm{d}\sigma(\tilde{\omega}) = \frac{|\det(M)|}{\|M\omega\|^3}\ \mathrm{d}\sigma(\omega).$$

So,

$$\begin{aligned}L(\tilde{\mathbf{x}}, \tilde{\omega})(\tilde{\omega} \cdot \tilde{\mathbf{n}})\ \mathrm{d}A(\tilde{\mathbf{x}})\mathrm{d}\sigma(\tilde{\omega}) &= L(\mathbf{x}, \omega)(\omega \cdot \mathbf{n})\ \mathrm{d}A(\mathbf{x})\mathrm{d}\sigma(\omega)\\ L(\tilde{\mathbf{x}}, \tilde{\omega})\left(\frac{M\omega}{\|M\omega\|} \cdot \frac{M^{-T}\mathbf{n}}{\|M^{-T}\mathbf{n}\|}\right)|\det(M)|\|M^{-T}\mathbf{n}\|\frac{|\det(M)|}{\|M\omega\|^3}\ \mathrm{d}A(\mathbf{x})\mathrm{d}\sigma(\omega) &= L(\mathbf{x}, \omega)(\omega \cdot \mathbf{n})\ \mathrm{d}A(\mathbf{x})\mathrm{d}\sigma(\omega)\\ L(\tilde{\mathbf{x}}, \tilde{\omega})(\omega \cdot \mathbf{n})\frac{|\det(M)|^2}{\|M\omega\|^4} &= L(\mathbf{x}, \omega)(\omega \cdot \mathbf{n})\\ L(\tilde{\mathbf{x}}, \tilde{\omega}) &= \frac{\|M\omega\|^4}{|\det(M)|^2}L(\tilde{\mathbf{x}}, \tilde{\omega}).\end{aligned}$$

Let us do some sanity check whether this satisfies the story we had before or not. Say, M is the uniform scaling by a factor of 2 in all direction. We have that $\|M\omega\| = 2\|\omega\| = 2$, and $\det(M) = 8$. So,

$$L(\tilde{\mathbf{x}}, \tilde{\omega}) = \frac{2^4}{8^2}L(\mathbf{x}, \omega) = \frac{16}{64}L(\mathbf{x}, \omega) = \frac{L(\mathbf{x}, \omega)}{4},$$

which matched our story above.

6 Appendix

Lemma 2.

$$\frac{\partial}{\partial u}\frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{1}{\|\mathbf{a}\|}\frac{\partial \mathbf{a}}{\partial u} - \frac{\mathbf{a}}{\|\mathbf{a}\|^3}\left(\mathbf{a} \cdot \frac{\partial \mathbf{a}}{\partial u}\right)$$

Proof.

$$
\begin{aligned}
\frac{\partial}{\partial u}\frac{\mathbf{a}}{\|\mathbf{a}\|} &= \frac{1}{\|\mathbf{a}\|^2}\left(\|\mathbf{a}\|\frac{\partial\mathbf{a}}{\partial u} - \mathbf{a}\frac{\partial\|\mathbf{a}\|}{\partial u}\right) = \frac{1}{\|\mathbf{a}\|^2}\left(\|\mathbf{a}\|\frac{\partial\mathbf{a}}{\partial u} - \mathbf{a}\frac{\partial\sqrt{\mathbf{a}\cdot\mathbf{a}}}{\partial u}\right) \\
&= \frac{1}{\|\mathbf{a}\|^2}\left(\|\mathbf{a}\|\frac{\partial\mathbf{a}}{\partial u} - \mathbf{a}\frac{1}{2\sqrt{\mathbf{a}\cdot\mathbf{a}}}\left(2\mathbf{a}\cdot\frac{\partial\mathbf{a}}{\partial u}\right)\right) \\
&= \frac{1}{\|\mathbf{a}\|^2}\left(\|\mathbf{a}\|\frac{\partial\mathbf{a}}{\partial u} - \frac{\mathbf{a}}{\|\mathbf{a}\|}\left(\mathbf{a}\cdot\frac{\partial\mathbf{a}}{\partial u}\right)\right) \\
&= \frac{1}{\|\mathbf{a}\|}\frac{\partial\mathbf{a}}{\partial u} - \frac{\mathbf{a}}{\|\mathbf{a}\|^3}\left(\mathbf{a}\cdot\frac{\partial\mathbf{a}}{\partial u}\right)
\end{aligned}
$$