Deriving the Smith shadowing function G_1 for $\gamma \in (0, 4]$

Rossen Dimov
rossen.dimov@chaosgroup.com

June 17, 2015 v1.1

Abstract

This document consists of detailed derivation of the equations used in the implementation of the GTR/GGX Microfacet reflection model in V-Ray 3.20

1 Table of symbols

D	Microfacet distribution function
G	Bidirectional shadowing-masking function
G_1	Monodirectional shadowing function
$\mathbf{i}$	Direction from which light is incident
$\mathbf{o}$	Direction in which light is scattered
$\mathbf{m}$	Microsurface normal
$\mathbf{n}$	Macrosurface normal
θ_m	Angle between $\mathbf{m}$ and $\mathbf{n}$
θ_v	Angle between $\mathbf{v}$ and $\mathbf{n}$, where $\mathbf{v}$ could be either $\mathbf{i}$ or $\mathbf{o}$
$\chi_+(a)$	Equal to 1 if $a > 0$ and zero if $a \leq 0$

2 Introduction

The Smith G approximates the bidirectional shadowing-masking as the separable product of two monodirectional shadowing terms G_1:

$$G(\mathbf{i}, \mathbf{o}, \mathbf{m}) \approx G_1(\mathbf{i}, \mathbf{m}) \, G_1(\mathbf{o}, \mathbf{m})$$

where G_1 is derived from the microfacet distribution D as described in Appendix A in [2]. For the given GGX distribution with roughness parameter α:

$$D(\mathbf{m}) = \frac{\alpha^2 \chi^+(\mathbf{m} \cdot \mathbf{n})}{\pi \cos^4 \theta_m (\alpha^2 + \tan^2 \theta_m)^2} \tag{33*}$$

they derive the following equation for G_1:

$$G_1(\mathbf{v},\mathbf{m}) = \chi^+\left(\frac{\mathbf{v}\cdot\mathbf{m}}{\mathbf{v}\cdot\mathbf{n}}\right)\frac{2}{1+\sqrt{1+\alpha^2\tan^2\theta_v}} \tag{34*}$$

Those equations are sufficient until control over the tail fall off of the reflection is required. This is enabled via Generalized-Trowbridge-Reitz, or GTR distribution defined in [1]:

$$D_{GTR}(\theta_h) = \frac{(\gamma-1)(\alpha^2-1)}{\pi(1-(\alpha^2)^{1-\gamma})}\frac{1}{(1+(\alpha^2-1)\cos^2\theta_h)^\gamma} \tag{1**}$$

Whenever we are using equations from the referenced papers we will stick to their original notation to avoid confusion. This is the reason the equations (33*) and (1**) are slightly different. However in our derivations below we are going to use just D_{GTR} without any parameters and θ without any subscripts.

In [1] are given the distribution $D_{GTR}(\theta_h)$ and the sampling equations for $\gamma \geq 0$. When $\gamma = 1$ there is a singularity and limit for $\gamma \to 1$ is used. It is also shown that when $\gamma = 2$ they get equivalent results to GGX distribution in [2]. However no derivations for G_1 for $\gamma \neq 2$ were made. In the next section we derive an approximation for computing G_1, which we measured to give less than 0.1% error compared to a numerical Monte Carlo solution.

3 G_1 approximation derivation

Given D_{GTR} our goal is to derive an equation for G_1 for $\gamma \geq 0$. Below is a list of the equations from [2] used in our derivation:

$$P_{22}(p,q) = D(\mathbf{m})\cos^4\theta_m \tag{42*}$$

$$\tan^2\theta_m = p^2+q^2 \tag{*}$$

$$P_2(q) = \int_{-\infty}^{\infty} P_{22}(p,q)\,\mathrm{d}p \tag{43*}$$

$$\mu = |\cot\theta_v| \tag{44*}$$

$$\Lambda(\mu) = \frac{1}{\mu}\int_{\mu}^{\infty}(q-\mu)\,P_2(q)\,\mathrm{d}q \tag{48*}$$

$$G_1(\mathbf{v},\mathbf{m}) = \chi^+\left(\frac{\mathbf{v}\cdot\mathbf{m}}{\mathbf{v}\cdot\mathbf{n}}\right)\frac{1}{1+\Lambda(\mu)} \tag{51*}$$

Given (1**) we plug it successively in (42*), (43*) and (48*):

$$P_{22}(p,q) = \frac{(\gamma-1)(\alpha^2-1)}{\pi(1-(\alpha^2)^{1-\gamma})}\frac{\cos^4\theta}{(1+(\alpha^2-1)\cos^2\theta)^\gamma} =$$

$$= \frac{(\gamma-1)(\alpha^2-1)}{\pi(1-(\alpha^2)^{1-\gamma})}\frac{\cos^4\theta}{(\sin^2\theta+\cos^2\theta\alpha^2)^\gamma} =$$

$$= \frac{(\gamma-1)(\alpha^2-1)}{\pi(1-(\alpha^2)^{1-\gamma})} \frac{(\cos^2\theta)^{2-\gamma}}{(\tan^2\theta+\alpha^2)^\gamma}$$

Using (*) gives $\cos^2\theta = \frac{1}{1+p^2+q2}$ and thus we get the following equation for P_{22}, P_2 and Λ:

$$P_{22}\left(p,q\right) = \frac{(\gamma-1)(\alpha^2-1)}{\pi(1-(\alpha^2)^{1-\gamma})} \frac{(p^2+q^2+1)^{\gamma-2}}{(p^2+q^2+\alpha^2)^\gamma}$$

$$P_2\left(q\right) = \frac{(\gamma-1)(\alpha^2-1)}{\pi(1-(\alpha^2)^{1-\gamma})} \int_{-\infty}^{\infty} \frac{(p^2+q^2+1)^{\gamma-2}}{(p^2+q^2+\alpha^2)^\gamma} \mathrm{d}p$$

$$\Lambda\left(\mu\right) = \frac{(\gamma-1)(\alpha^2-1)}{\mu\pi(1-(\alpha^2)^{1-\gamma})} \int_{\mu}^{\infty} \int_{-\infty}^{\infty} \frac{(q-\mu)(p^2+q^2+1)^{\gamma-2}}{(p^2+q^2+\alpha^2)^\gamma} \mathrm{d}p\mathrm{d}q \qquad (1)$$

Unfortunately we couldn't derive solution in closed form for Λ for an arbitrary γ. So we decided we

- try to derive closed form solution for a certain values of γ
- plot all G_1-s as height fields depending on μ and α
- based on the visual results, judge if it's worth interpolating for intermediate values of γ
- pick an interpolating method based on rendering results
- estimate the error between interpolation and Monte Carlo numerical solution

Using Mathematica we derived closed form solution for $\gamma = 0$, 1, 2, 3 and 4. For Mathematica code see files G1.nb and G1_1.nb.

3.1 Deriving G_1^0

For $\gamma = 0$:

$$P_{22}^0\left(p,q\right) = \frac{1}{\pi(p^2+q^2+1)^2}$$

$$P_2^0\left(q\right) = \frac{1}{2(q^2+1)^{3/2}}$$

$$\Lambda^0\left(\mu\right) = \frac{\sqrt{\mu^2+1}-\mu}{2\mu}$$

$$G_1^0\left(\mu\right) = \frac{2\mu}{\sqrt{\mu^2+1}+\mu} = \frac{2}{1+\sqrt{\frac{1}{\mu^2}+1}}$$

3.2 Deriving G_1^1

For $\gamma = 1$ we used

$$D_{GTR_1}(\theta_h) = \frac{(\alpha^2 - 1)}{\pi \log \alpha^2} \frac{1}{(1 + (\alpha^2 - 1)\cos^2 \theta_h)} \tag{4**}$$

shown in [1].

$$P_{22}^1(p, q) = \frac{\alpha^2 - 1}{\pi(p^2 + q^2 + 1)(p^2 + q^2 + \alpha^2)\log(\alpha^2)}$$

$$P_2^1(q) = \frac{1}{\log(\alpha^2)} \left(\frac{1}{\sqrt{q^2 + 1}} - \frac{1}{\sqrt{q^2 + \alpha^2}} \right)$$

$$\Lambda^1(\mu) = \frac{\sqrt{\mu^2 + \alpha^2} - \sqrt{\mu^2 + 1} + \mu \operatorname{arcsinh}(\mu) - \mu \log(\mu + \sqrt{\mu^2 + \alpha^2})}{\mu \log \alpha^2}$$

$$G_1^1(\mu) = \frac{\mu \log \alpha^2}{\sqrt{\mu^2 + \alpha^2} - \sqrt{\mu^2 + 1} + \mu \operatorname{arcsinh}(\mu) + \mu \log \alpha^2 - \mu \log(\mu + \sqrt{\mu^2 + \alpha^2})}$$

In the implementation the term for G_1^1 has been further simplified using $\operatorname{arcsinh}(z) = \log(z + \sqrt{z^2 + 1})$, logarithmic identities product and quotient, and setting $A = \sqrt{\mu^2 + \alpha^2}$ and $B = \sqrt{\mu^2 + 1}$:

$$G_1^1(\mu) = \frac{\mu \log \alpha^2}{A - B + \mu \log\left(\frac{\alpha^2(\mu + B)}{\mu + A}\right)}$$

3.3 Deriving G_1^2

For $\gamma = 2$ the derivation was done only to see if we will get the same results as in [2]:

$$P_{22}^2(p, q) = \frac{\alpha^2}{\pi(p^2 + q^2 + \alpha^2)^2}$$

$$P_2^2(q) = \frac{\alpha^2}{2(q^2 + \alpha^2)^{3/2}}$$

$$\Lambda^2(\mu) = \frac{\sqrt{\mu^2 + \alpha^2} - \mu}{2\mu}$$

$$G_1^2(\mu) = \frac{2\mu}{\sqrt{\mu^2 + \alpha^2} + \mu} = \frac{2}{1 + \sqrt{\frac{\alpha^2}{\mu^2} + 1}} = (34^*)$$

3.4 Deriving G_1^3

For $\gamma = 3$:

$$P_{22}^3(p,q) = \frac{2\alpha^4(p^2+q^2+1)}{\pi(\alpha^2+1)(p^2+q^2+\alpha^2)^3}$$

$$P_2^3(q) = \frac{\alpha^4(\alpha^2+4q^2+3)}{4(\alpha^2+1)(q^2+\alpha^2)^{5/2}}$$

$$\Lambda^3(\mu) = \frac{3\alpha^4+2\mu(\mu-\sqrt{\alpha^2+\mu^2})+\alpha^2\left(1+2\mu\left(\mu-\sqrt{\alpha^2+\mu^2}\right)\right)}{4(1+\alpha^2)\mu\sqrt{\alpha^2+\mu^2}}$$

$$G_1^3(\mu) = \frac{4(1+\alpha^2)\mu\sqrt{\alpha^2+\mu^2}}{3\alpha^4+2\mu(\mu+\sqrt{\alpha^2+\mu^2})+\alpha^2\left(1+2\mu\left(\mu+\sqrt{\alpha^2+\mu^2}\right)\right)}$$

In the implementation the term for G_1^3 has been further simplified by regrouping and setting $A = \sqrt{\mu^2+\alpha^2}$ and $B = \alpha^2+1$:

$$G_1^3(\mu) = \frac{4B\mu A}{\alpha^2(3\alpha^2+1)+2\mu B(\mu+A)}$$

3.5 Deriving G_1^4

For $\gamma = 4$:

$$P_{22}^4(p,q) = \frac{3\alpha^6(p^2+q^2+1)^2}{\pi(\alpha^4+\alpha^2+1)(p^2+q^2+\alpha^2)^4}$$

$$P_2^4(q) = \frac{3\alpha^6(8q^4+4\alpha^2q2+12q^2+\alpha^4+2\alpha^2+5)}{16(\alpha^4+\alpha^2+1)(q^2+\alpha^2)^{7/2}}$$

$$\Lambda^4(\mu) = \frac{15\alpha^8+3\alpha^4+8\mu^3(\mu-\sqrt{\alpha^2+\mu^2})+\alpha^6(6+24\mu^2-8\mu\sqrt{\alpha^2+\mu^2})}{16(\alpha^4+\alpha^2+1)\mu(\mu^2+\alpha^2)^{3/2}}+$$
$$+\frac{4\alpha^2\mu(-2\sqrt{\alpha^2+\mu^2}+\mu(3+2\mu(\mu-\sqrt{\alpha^2+\mu^2})))(\alpha^2+1)}{16(\alpha^4+\alpha^2+1)\mu(\mu^2+\alpha^2)^{3/2}}$$

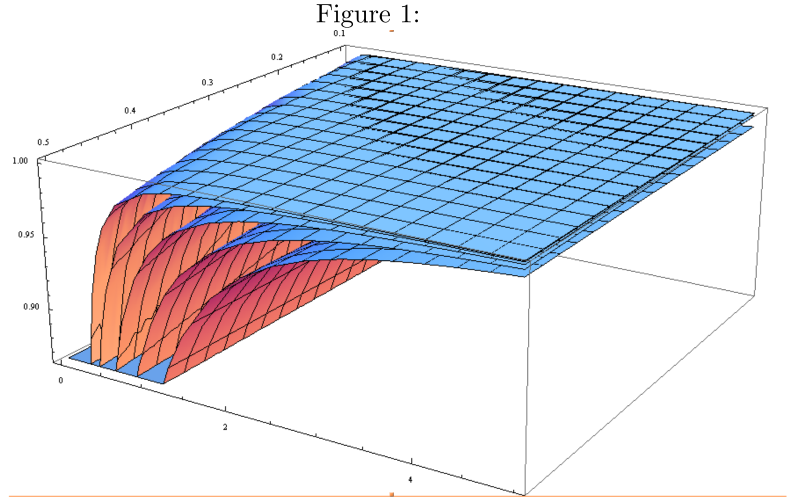
In the implementation the term for G_1^4 has been simplified by regrouping and setting $A = 8\alpha^4+8\alpha^2+8$ and $B = \sqrt{\mu^2+\alpha^2}$:

$$G_1^4(\mu) = \frac{2A\mu B^3}{A\mu(B^3+\mu^3)+3\alpha^2(\alpha^2(5\alpha^4+2\alpha^2+1)+4\mu^2(2\alpha^4+\alpha^2+1))}$$

4 Plotting $G_1^0, G_1^1, G_1^2, G_1^3$, and G_1^4

Using the derived equations in Section 3 we used Mathematica to plot them. For the plot shown here we set $\alpha \in (0.1, 0.5)$ and $\mu \in (0, 5)$:

Figure 1:



All height fields converge to G_1^0 when $\mu \to 1$ and $\alpha \to 0$.

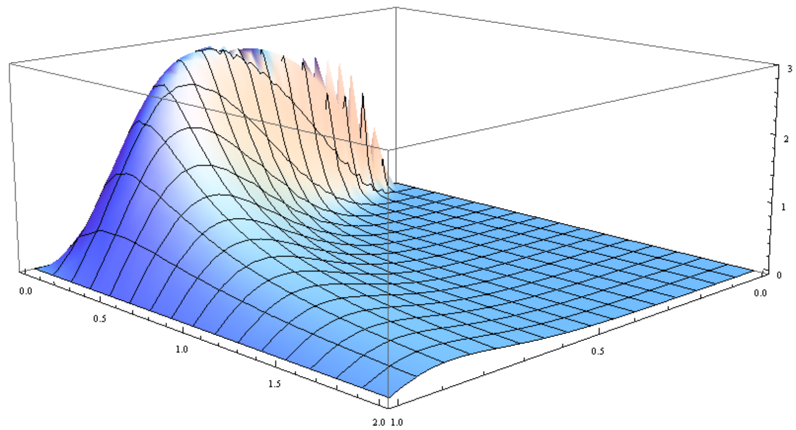
5 Interpolation methods and error

First we tried linear interpolation but we got bad looking results (especially for γ slightly less than 1), where ‘bad looking results’ mean quite a big visual difference for very little change of γ. Then we tried spline interpolation, we were satisfied with the results and this is what we finally used in our implementation. We finally compared the interpolated results with numerical Monte Carlo results and we get ca. 0.1% difference - G_1 returns values between $(0, 1)$ and the mean difference was 0.001.

6 $\gamma > 4$

In the current implementation γ is clamped up to 4. Initially we decided we can ignore higher values since the rendered results seemed fine and we couldn’t derive equations for $\gamma > 4$. Moreover values greater than 10 are impractical for the GTR/GGX. However later we succeeded and bellow is a plot for $\left|G_1^4 - G_1^{10}\right|$ multiplied 10x with $\alpha \in (0, 1)$ and $\mu \in (0, 2)$:

Figure 2:



From the plot we can observe that the biggest difference is in $\mu \in (0,1)$ and it s up to 0.3. It is not as negligible as we thought but since we are happy with current rendering results and the equation for G_1^{10} is quite large we decided to leave it as it is.

7 Remapping the roughness α in the implementation

In the current implementation the roughness α is remapped with $(1-\alpha)^2$ and is named *sharpness*.

References

[1] Brent Burley. Physically based shading at disney, 2012.

[2] Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics Conference on Rendering Techniques*, EGSR'07, pages 195–206, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.