The Comprehensive PBR Guide by Allegorithmic - vol. 1

Light and Matter : The theory of Physically-Based Rendering and Shading

Cover by Gaëtan Lassagne, written by Wes McDermott

allegorithmic
innovation matters

Table of Contents

Technical edit by: Cyrille Damez and Nicolas Wirrmann

Light and Matter

The theory of Physically-Based Rendering and Shading

Light is a complex phenomenon as it can exhibit properties of both a wave and a particle. As a result, different models have been created to describe the behavior of light. As texture artists, we are interested in the Light Ray Model as it describes the interaction of light and matter. It's important for us to understand how light rays interact with surface matter because our job is to create textures that describe a surface. The textures and materials we author interact with light in our virtual worlds and the more we understand about how light behaves, the better our textures will look.

In this guide, we will discuss the theory behind the physics through which physically-based rendering models are based upon. We will start with a light ray and work up to defining the key factors for PBR.

Light Rays

The Light Ray Model states that a light ray has the trajectory of a straight line in homogeneous transparent media such as air. The Light Ray Model also says that the ray will behave in a predictable manner when encountering surfaces such as opaque objects or passing through a different medium such as air to water. This makes it possible to visualize the path the light ray will follow as it moves from a starting point to where it eventually changes into another form of energy such as heat.

The light ray that hits a surface is called the Incident Ray and the angle that at which it hits is called the Angle of Incidence as shown in figure 01.

A light ray is incident on a plane interface between two media.

When a light ray hits a surface, either or possibly both of these things can happen:

1. The light ray is reflected off the surface and travels in a different direction. It follows the Law of Reflection, which states that the Angle of Reflection is equal to the Angle of Incidence (Reflected Light).

2. The light ray passes from one medium to another in the trajectory of a straight line (Refracted Light).

At this point, we can state that light rays split into two directions: reflection and refraction. At the surface, the ray is either reflected or refracted and it can be eventually absorbed by either medium. However, absorption doesn't occur at the surface.

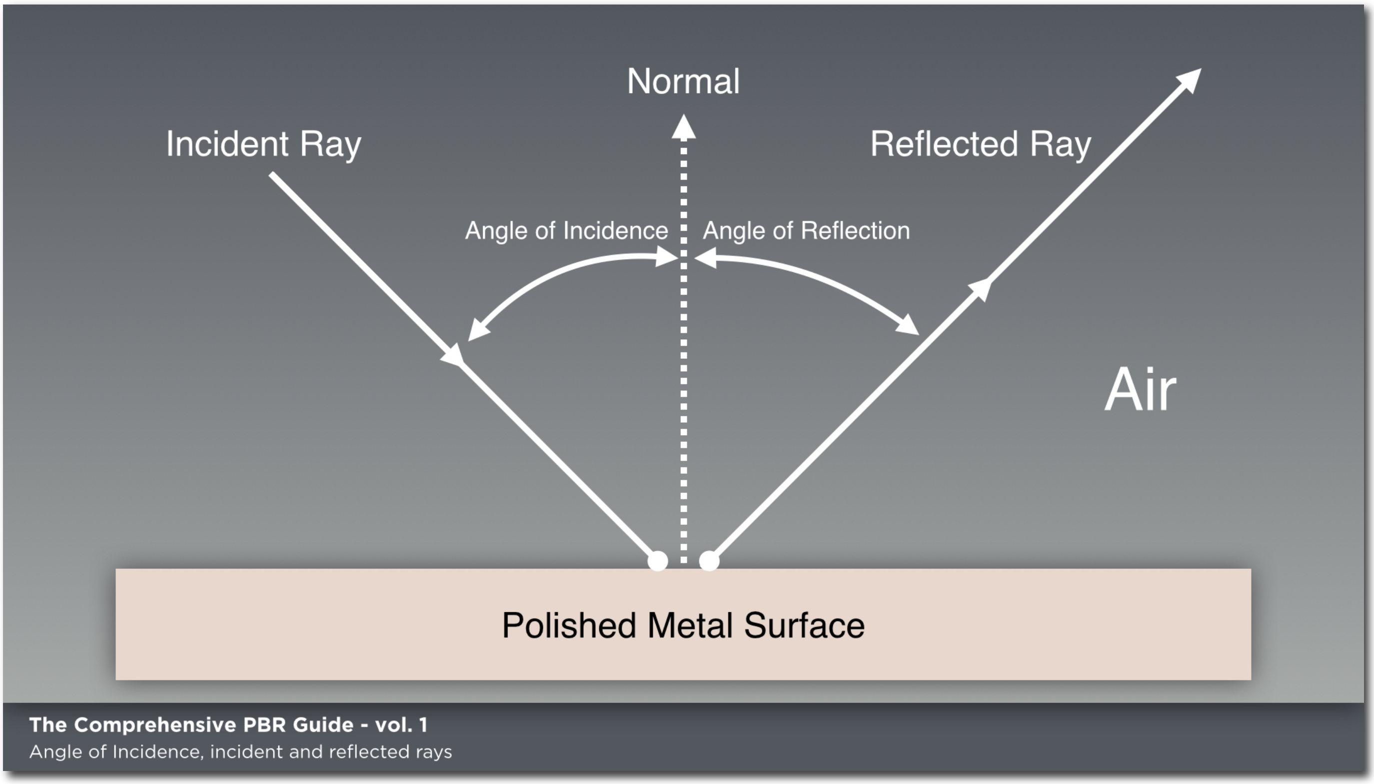


Figure 01

Absorption and Scattering (Transparency and Translucency)

When traveling in an inhomogeneous medium or translucent material, light can be absorbed or scattered:

1. With absorption, the light intensity decreases as it is changed into another form of energy (usually heat), and its color changes as the amount of light absorbed depends on the wavelength, but the direction of the ray doesn't change.

2. With scattering, the ray direction is changed randomly, the amount of deviation depending on the material. Scattering randomizes light direction but the intensity doesn't change. An ear is a good example. The ear is thin (absorption is low), so you can see the scattered light penetrating out of the back of the ear. If there is no scattering and the absorption is low, rays can pass directly through the surface such as with glass. For example, if you are swimming in a pool, which is hopefully clean, you can open your eyes and see at a fairly good distance through the clear water. However, let's imagine that same pool hasn't been cleaned in a while and the water is dirty. The dirt particles scatter the light and thus make the clarity of the water much lower.

The further light travels in such a medium/material, the more it is absorbed and/or scattered. Therefore, object thickness plays a large role in how much the light is absorbed or scattered. A thickness map can be used to describe object thickness to the shader as shown in figure 02.

Object thickness plays a large role in how much the light is absorbed or scattered

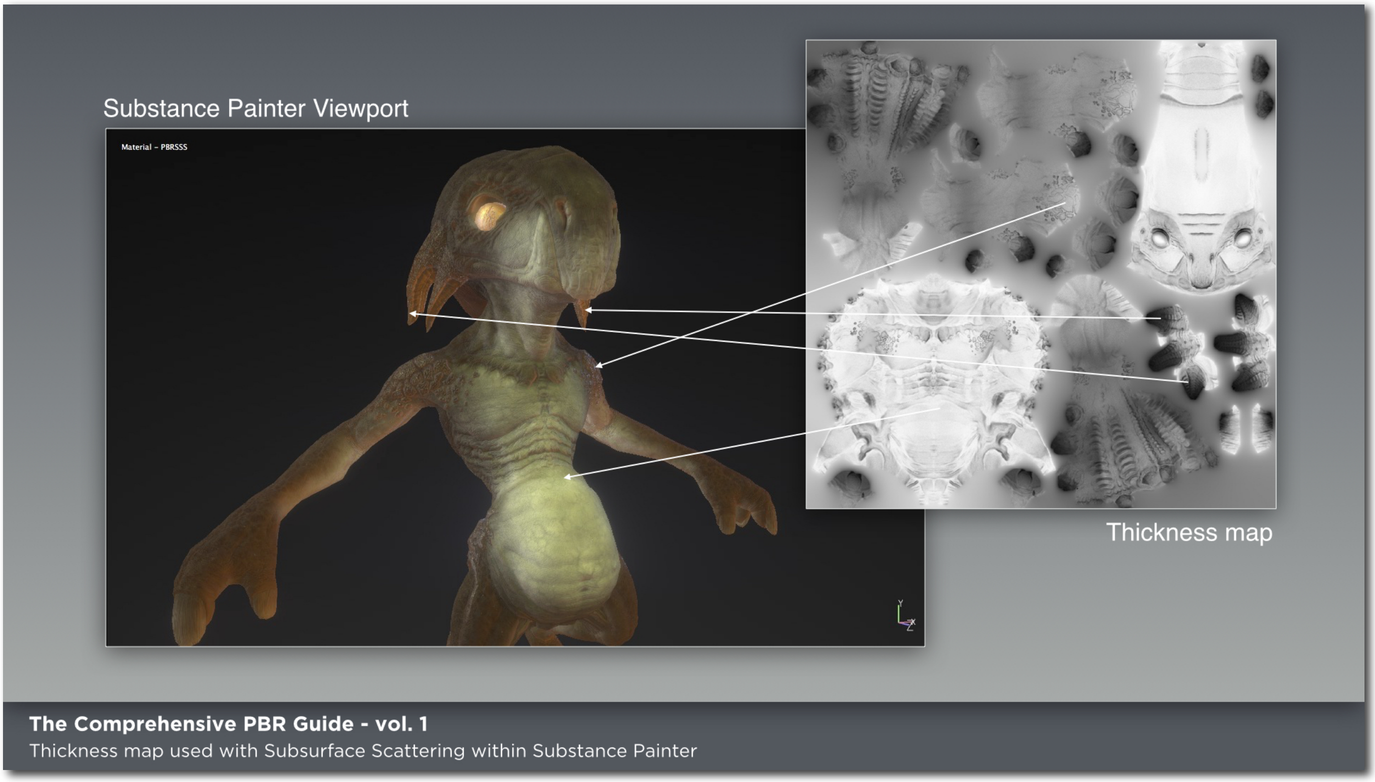


Figure 02

Diffuse and Specular Reflection

Specular reflection is light that has been reflected at the surface, as we discussed above in the Light Ray section. The light ray is reflected off the surface and travels in a different direction. It follows the Law of Reflection, which states that on a perfectly planar surface the Angle of Reflection is equal to the Angle of Incidence. However, it is important to note that most surfaces are irregular and that the reflected direction will therefore vary randomly based on the surface roughness. This changes light direction, but the light intensity remains constant.

Rougher surfaces will have larger and dimmer looking highlights. Smoother surfaces will keep specular reflections focused, which can appear to look brighter or more intense when looked at from the proper angle. However, the same total amount of light is reflected in both cases as shown in figure 03.

Diffuse reflection is light that has been refracted. The light ray passes from one medium to another and is scattered multiple times inside the object. Then it is refracted again out of the object making its way back to the original medium at approximately the same point where it went through the first time as shown in figure 04.

Diffuse materials are fairly absorbent, meaning that if the refracted light travels for too long in that material, it has a good chance of being completely absorbed. This means that if the light ever comes out of that material, it has probably not traveled very far from the point of entry. That's why the distance between the entry and exit points can be neglected. The Lambertian model, which is usually used for diffuse reflection in a traditional shading sense, does not take surface roughness into account, but there are diffuse reflection models that do such as Oren-Nayar.

Materials that have both high scattering but low absorption are sometimes referred to as "participating media" or "translucent materials". Examples of these are smoke, milk, skin, jade and marble. Rendering of the latter three may be possible with the additional modeling of subsurface scattering where the difference between the ingoing and outgoing point of the light ray is no longer neglected. Accurate rendering of medium with highly varying and very low scattering and absorption like smoke or fog may require even more expensive methods such as Monte Carlo simulations.

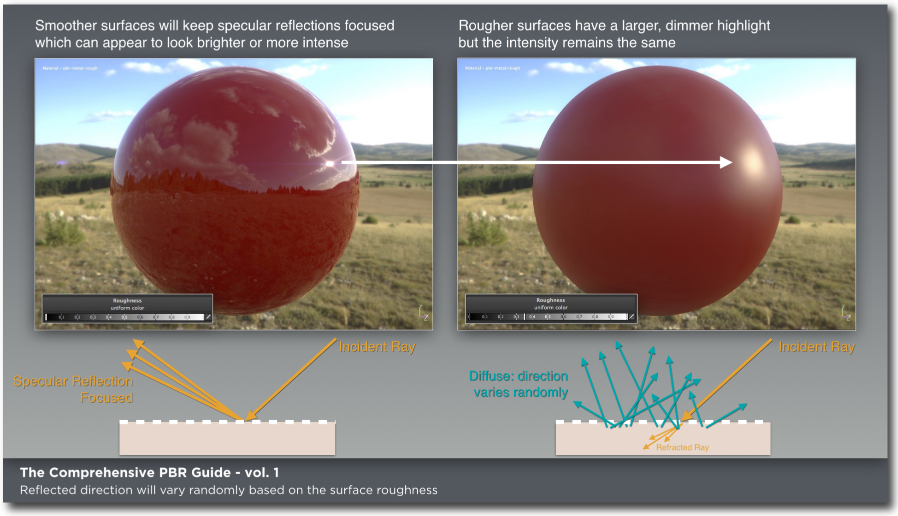


Figure 03

Rougher surfaces will have larger and dimmer looking highlights

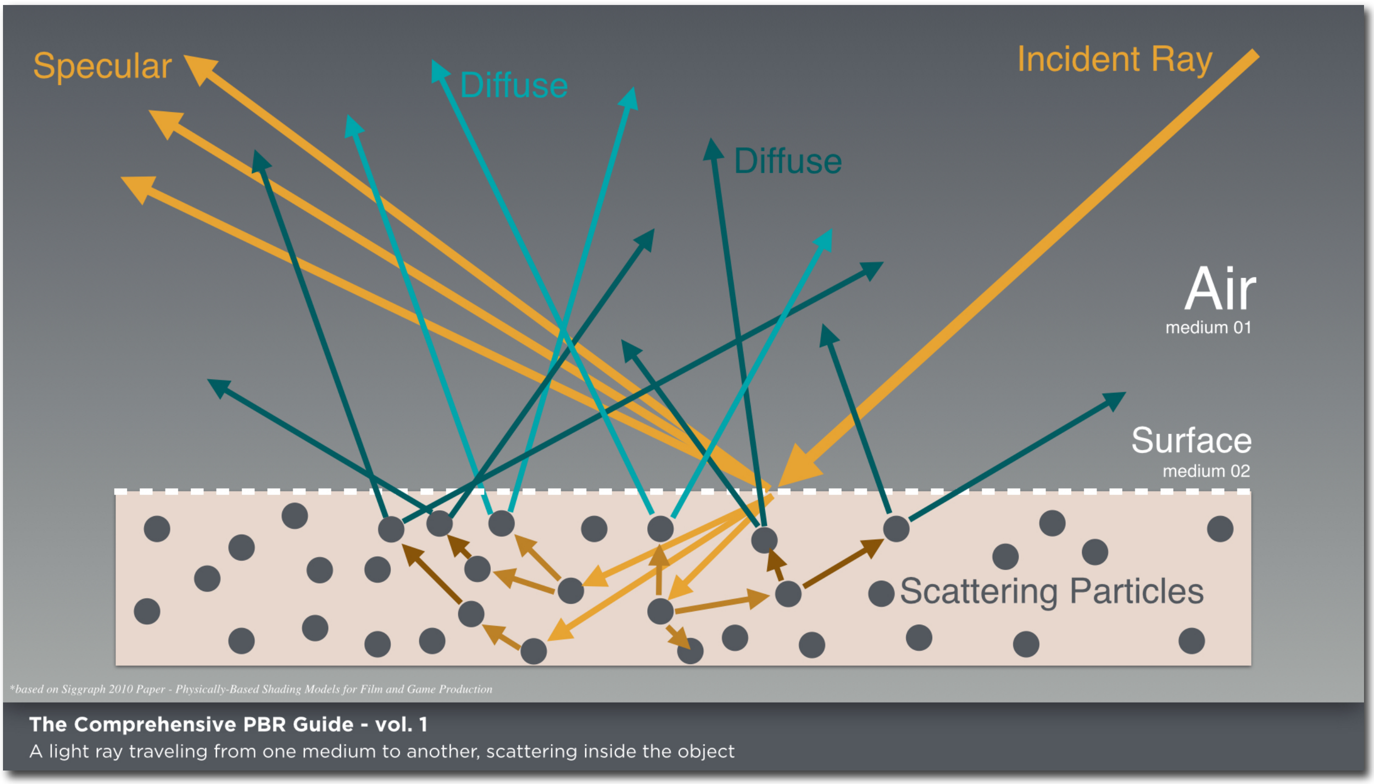
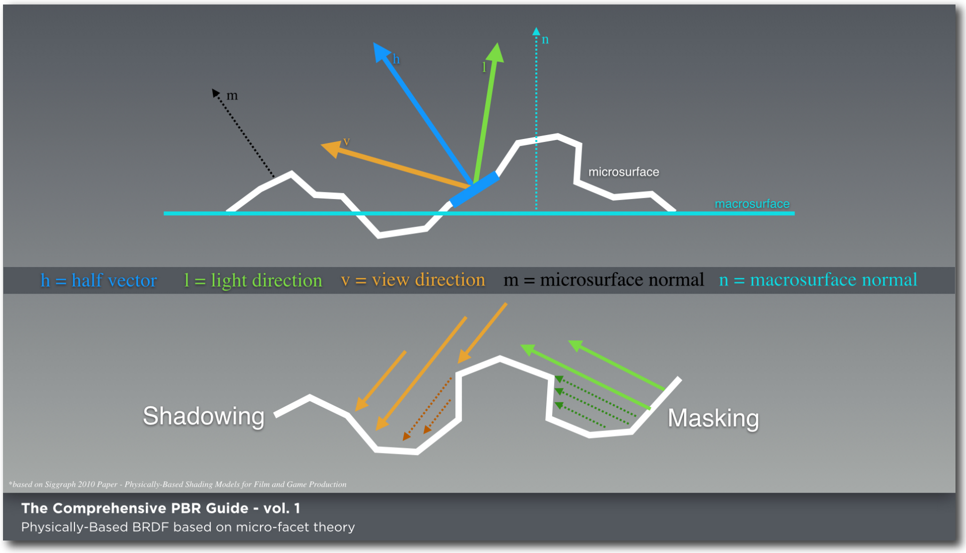


Figure 04

Microfacet Theory

In theory, both diffuse and specular reflection are dependent on the surface irregularities where the light rays intersect. In practice though, the effect of roughness on diffuse reflection is much less visible because of the scattering happening inside the material. As a result, the outgoing direction of the ray is fairly independent of surface roughness and the incident direction. The most common model for diffuse reflection (Lambertian) completely neglects it.



The Comprehensive PBR Guide - vol. 1
Physically-Based BRDF based on micro-facet theory

Figure 05

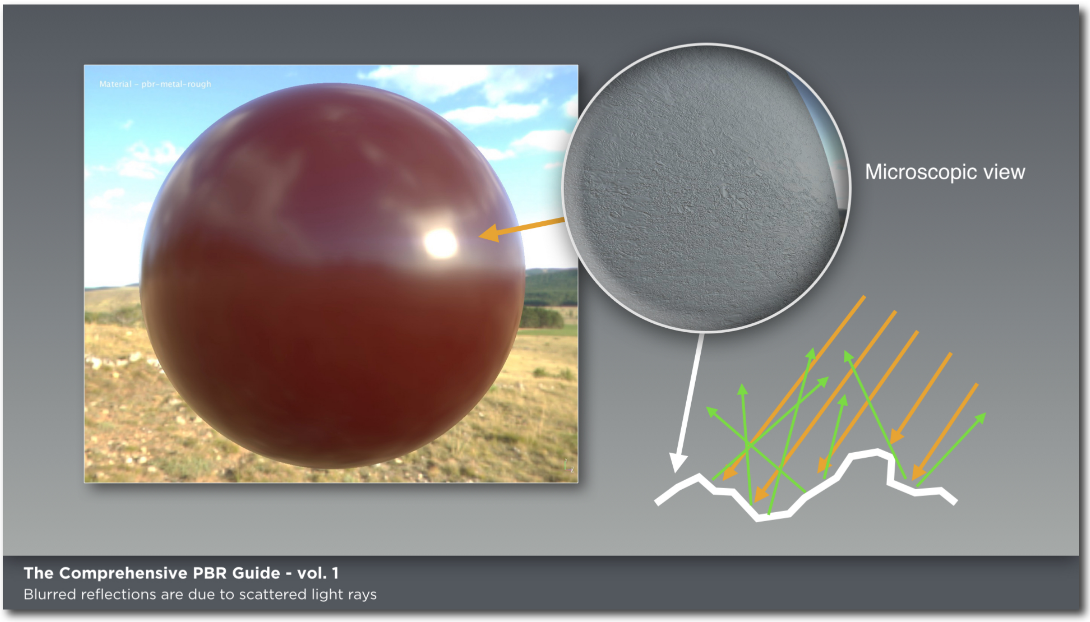
In this document, we have referred to these surface irregularities as surface roughness. Actually, it is often referred to by several names such as roughness, smoothness, glossiness or micro-surface, depending on the PBR workflow in use, but they describe the same aspect of a surface, which is sub-texel geometric detail.

These surface irregularities are authored in the roughness or glossiness map depending on the workflow you are using. A physically-based BRDF is based on the microfacet theory which supposes that a surface is composed of small-scaled planar detail surfaces of varying orientation called microfacets. Each of these small planes reflects light in a single direction based on its normal as shown in figure 05.

Micro-facets whose surface normal is oriented exactly halfway between the light direction and view direction will reflect visible light. However, not all microfacets where the microsurface normal and the half normal are equal will contribute as some will be blocked by shadowing (light direction) or masking (view direction) as is illustrated in figure 05.

The surface irregularities at a microscopic level cause light diffusion. For example, blurred reflections are due to scattered light rays. The rays are not reflected in parallel so we perceive the specular reflection as blurred as shown in figure 06.

The surface irregularities at a microscopic level cause light diffusion



The Comprehensive PBR Guide - vol. 1
Blurred reflections are due to scattered light rays

Figure 06

Color

The color of a surface (which is to say the color that we see) is due to which wavelengths are emitted by the light source, which are absorbed by the object and which others are reflected both specularly and diffusely. The remaining reflected wavelengths are what we see as color.

For example, the skin of an apple mostly reflects red light. Only the red wavelengths are scattered back outside the apple skin and the others are absorbed by it as shown in figure 07.

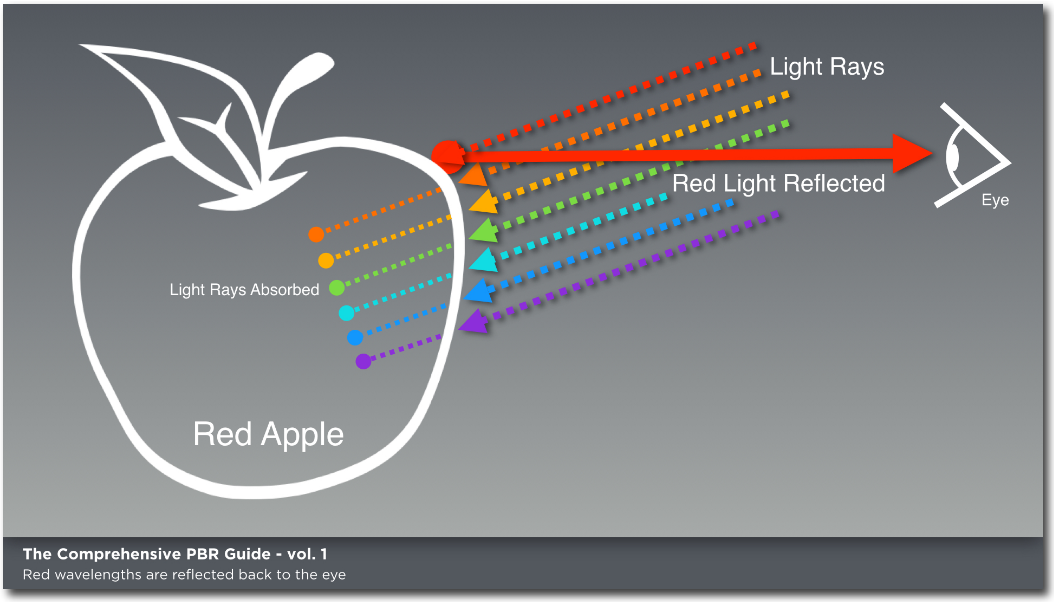


Figure 07

It also has bright specular highlights the same color as the light source because with materials like the skin of an apple that are not electrical conductors (dielectrics), specular reflection is almost independent of wavelength. Therefore, for such materials the specular reflection is never colored. We will discuss more about the different type of materials (metals and dielectrics) in later sections.

Substance PBR shaders use the GGX microfacet distribution

BRDF

A Bidirectional Reflectance Distribution Function (BRDF) simply put is a function that describes the reflectance properties of a surface. In computer graphics, there are different BRDF models some of which are not physically plausible. For a BRDF to be physically plausible, it must be energy conserving and exhibit reciprocity. For reciprocity, I am referring to the Helmholtz Reciprocity principle, which states that incoming and outgoing light rays can be considered as reversals of each other without affecting the outcome of the BRDF.

The BRDF used by Substance's PBR shaders is based on Disney's "principled" reflectance model, which is based on the GGX microfacet distribution. GGX provides one of the better solutions in terms of specular distribution in that it has a shorter peak in the highlight and a longer tail in the falloff, which is to say that it looks more realistic as shown in figure 08.

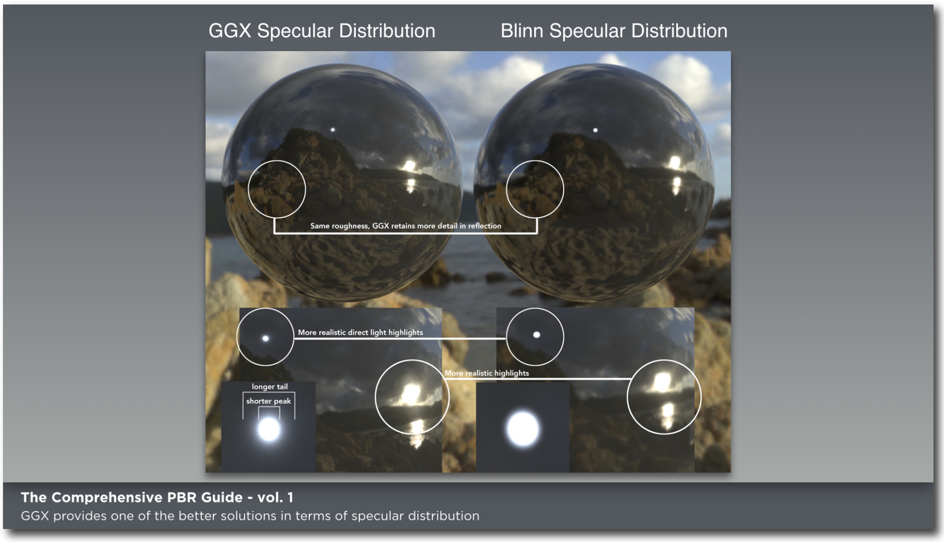


Figure 08

Object GGX provides one of the better solutions in terms of specular distribution

Energy Conservation

Energy Conservation plays a vital role in physically-based rendering solutions. It states that the total amount of light re-emitted by a surface (reflected and scattered back) is less than the total amount it received. In other words, the light reflected off the surface will never be more intense than it was before it hit the surface. As artists, we don't have to worry about controlling Energy Conservation. This is one of the nice aspects of PBR in that energy conservation is always enforced by the shader. It’s part of the physically-based model and it allows us to focus more on art rather than physics.

Fresnel Effect

The Fresnel reflection factor also plays a vital role in physically-based shading as a coefficient of the BRDF. The Fresnel Effect as observed by French physicist Augustin-Jean Fresnel states that the amount of light you see reflected from a surface depends on the viewing angle at which you perceive it.

For example, think of a pool of water. If you look straight down, perpendicular to the water surface, you can see down to the bottom. Viewing the water surface in this manner would be at zero degrees or normal incidence, normal being the surface normal. Now, if you look at the pool of water at a grazing incidence, more parallel to the water surface, you will see that the specular reflections on the water surface become more intense and you may not be able to see below the surface of the water at all.

Fresnel is not something that we control in PBR as we did in traditional shading. Again, this is another physics aspect that is handled for us by the PBR shader. When it comes to viewing a surface at a grazing incidence, all smoothed surfaces will become a nearly 100% reflector at a 90 degree angle of incidence.

For rough surfaces, reflectance will become increasingly specular but we won't approach 100% specular reflection. What matters then is the angle between the normal of each microfacet and the light, not the angle between the normal of the "macrosurface" and the light. Because the light rays are dispersed into different directions, the reflection appears softer or dimmer. What you get at a macroscopic level is a bit like the average of all the Fresnel effect you would have for the microfacets.

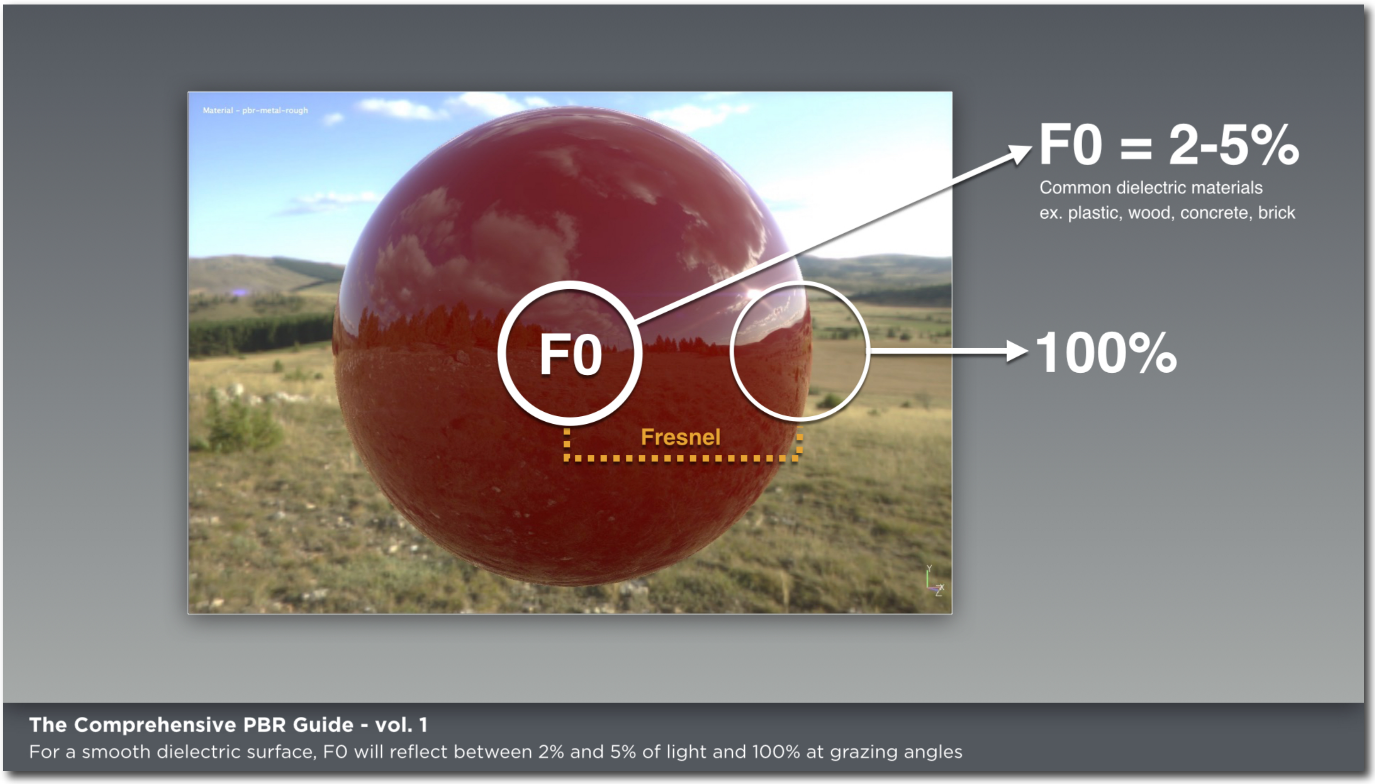


Figure 09

For rough surfaces, reflectance will become increasingly specular but we won't approach 100% specular reflection

F0 (Fresnel Reflectance at 0 Degrees)

When light hits a surface straight on or perpendicularly (0 degree angle), there is a percentage of that light that is reflected back as specular. Using the Index of Refraction (IOR) for a surface, you can derive the amount that is reflected back and this is referred to as F0 (Fresnel 0) as shown in figure 09. The amount of light that is refracted into the surface is referred to a 1-F0.

The F0 range for most common dielectrics will be from 0.02 - 0.05 and for conductors the F0 range will be 0.5-1.0. Thus, the reflectivity of a surface is determined by the refractive index as shown in the following equation from Sebastien Lagarde's "Feeding a Physically-based Shading Model" blog post as shown in figure 10.

$$F(0^\circ) = \frac{(n-1)^2}{(n+1)^2} = 0.02$$

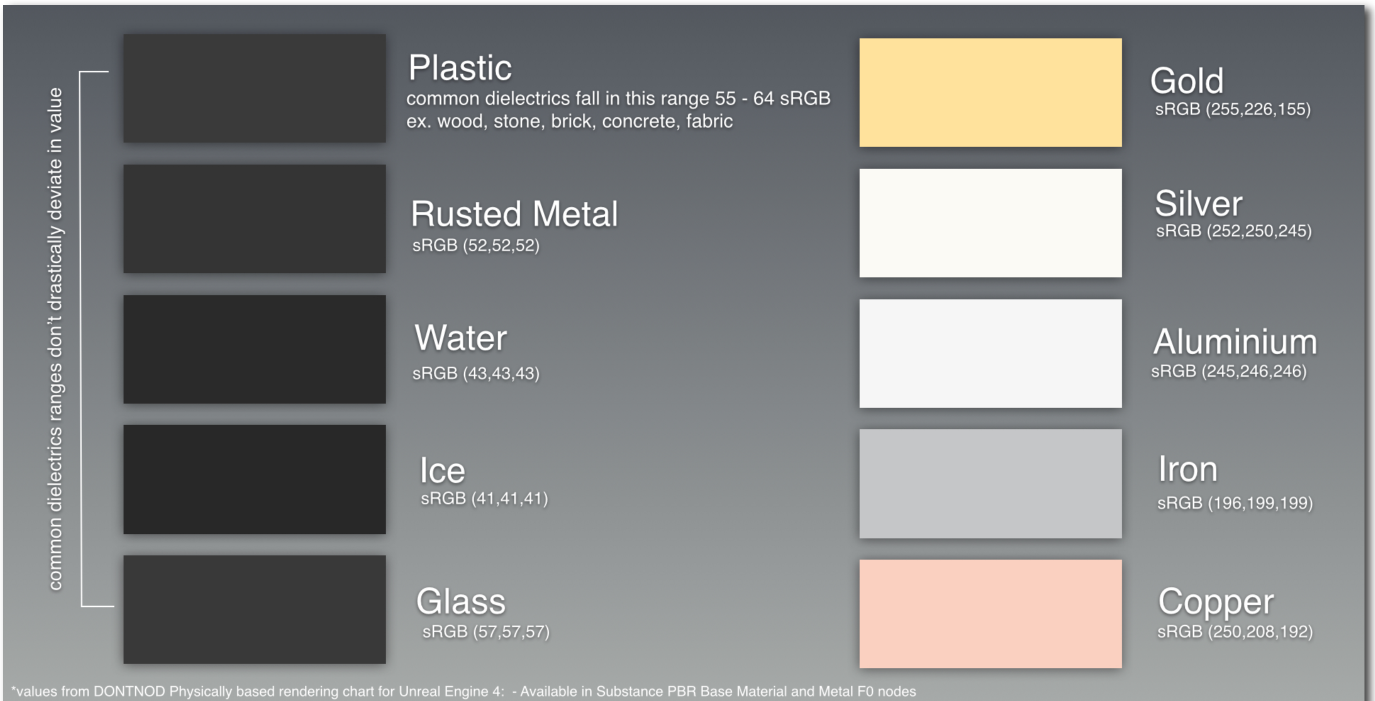
The Comprehensive PBR Guide - vol. 1
Reflectivity of a surface is determined by the refractive index

Figure 10

It is the F0 reflectance value that we are concerned with in regards to authoring our textures. Non-metals (dielectrics/insulators) will have a greyscale value and metals (conductors) will have an RGB value. With regards to PBR and from an artistic interpretation of reflectance, we can state that for a common smooth dielectric surface, F0 will reflect between 2% and 5% of light and 100% at grazing angles as was shown in figure 09.

The dielectric (non-metal) reflectance values don't actually change very drastically. In fact, when altered by roughness the actual changes in value can be hard to see. However, there is a difference in the values. In figure 11, you can see a chart that shows the F0 ranges for both metal and non-metal materials.

Notice that the ranges for non-metals do not deviate from each other drastically. Gemstones are an exception as they have higher values. We will discuss F0 as it specifically relates to conductors and insulators a bit later.



The Comprehensive PBR Guide - vol. 1
F0 ranges for both metal and non-metal materials

Figure 11

Conductors and Insulators (Metals and Non-Metals)

When creating materials for PBR, I find it helpful to think in terms of metal or non-metal. I simply ask myself if the surface is metal or not. If it is, I follow one set of guidelines and if it's not, I follow another. This can be a rather simplistic approach as some materials may not fall into these categories such as metalloids, but in the overall process of creating materials, distinguishing between metal and non-metal is a good approach and metalloids are an exception. To set up guidelines for materials, we first must understand what we are trying to create. With PBR, we can look at the properties of metals (conductors) and non-metals (insulators) to derive this set of guidelines.

Refracted light is absorbed, the color tint of metals come from the reflected light and thus in our maps, we don't give metals a diffuse color

Metals

Metals (conductors) are good conductors of heat and electricity. Simply put, the electric field in conducting metals is zero and when an incoming light wave made of electric and magnetic fields hits the surface, it is partially reflected and all the refracted light is absorbed. The reflectance value for polished metal is going to be high at a range of about 70-100% reflective as shown in figure 12.

Some metals absorb light at different wavelengths. For example, gold absorbs blue light at the high-frequency end of the visible spectrum so it appears yellow as a result. However, since the refracted light is absorbed, the color tint of metals come from the reflected light and thus in our maps, we don't give metals a diffuse color. For example, in the specular/gloss workflow, raw metal is set to black in the diffuse map and the reflectance value is a tinted color value in the specular map. With metals, the reflectance value will be RGB and can be tinted. Since we are working within a physically-based model, we need to use real-world measured values for the metal reflectance in our maps.

Another important aspect with metals in terms of texturing is that metal can corrode. This means that weathering elements can play a large role in the reflective state of metal. If the metal rusts for example, this changes the reflective state of the metal and the corroded areas are then treated as a dielectric material as shown in figure 13.

Also, metal that is painted is not treated like a metal but rather a dielectric as well. The paint acts as a layer on top of the raw metal. Only the raw metal exposed from chipped away paint is treated as metal. The same goes for dirt on the metal or any matter that obscures the raw metal.

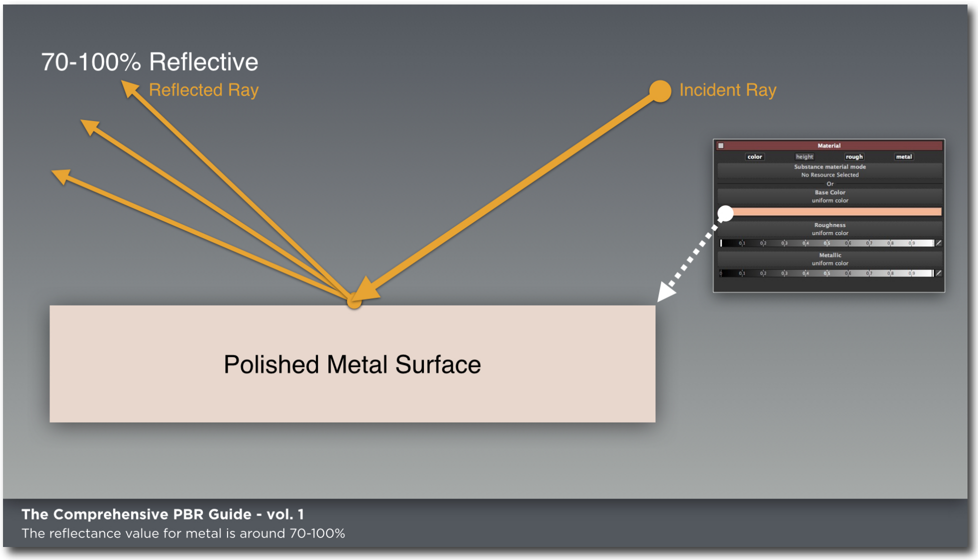


Figure 12

I stated above that I always ask myself if a material is a metal or not. However, to be more precise, the question should also inquire the state of the metal such as is it painted, rusted or covered in dirt/grease. The material will be treated as dielectric if it is not raw metal and there could be some blending between metal and non-metal depending on the weathering.

Weathering elements can play a large role in the reflective state of metal

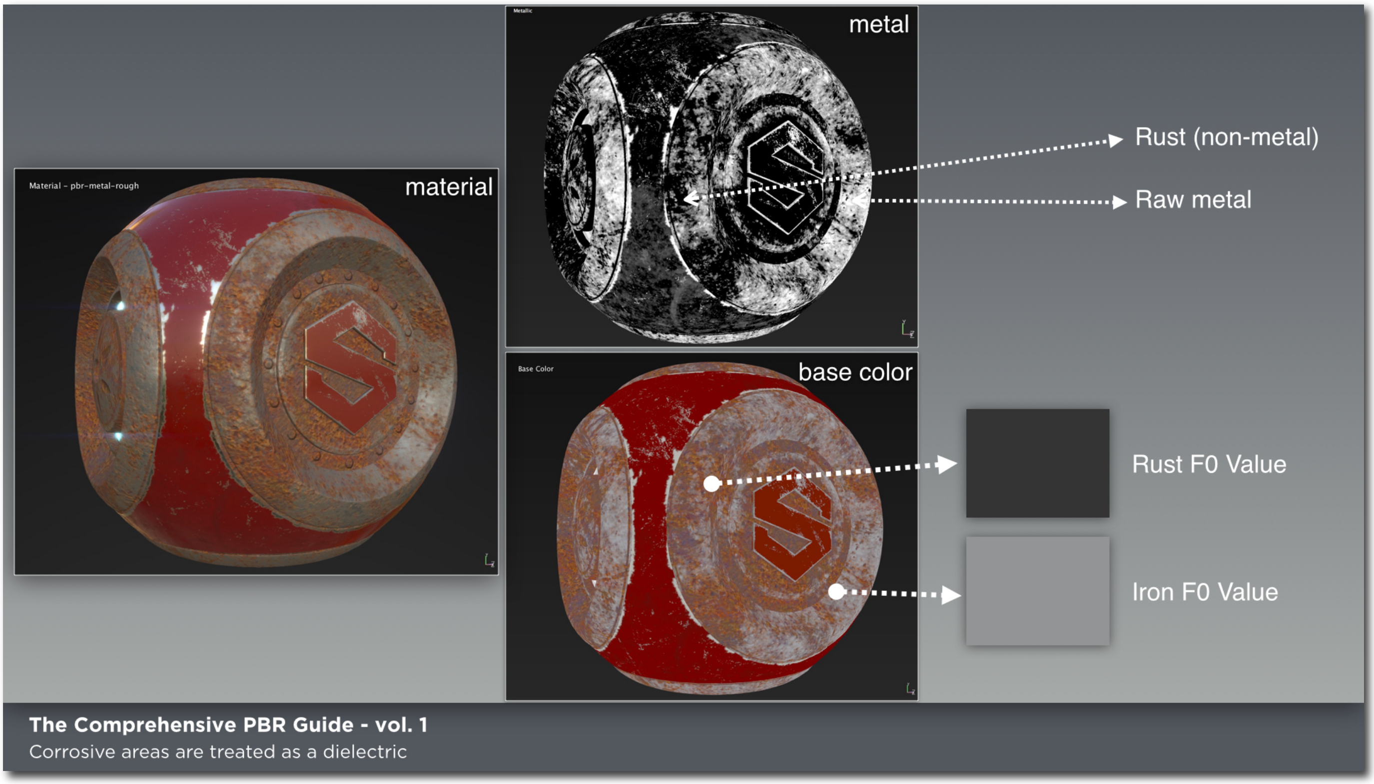


Figure 13

Non-Metals

Non-metals (insulators/dielectrics) are poor conductors of electricity. The refracted light is scattered and/or absorbed (often re-emerging from the surface) and thus they reflect a much smaller amount of light than metals and will have an albedo color. We stated earlier that the value for common dielectrics would be around 2-5% based on the F0 as computed by the index of refraction. These values are contained within the linear range of 0.017-0.067 (40-75 sRGB) as shown in figure 14. With the exception of gemstones, most dielectrics will not be greater than 4%.

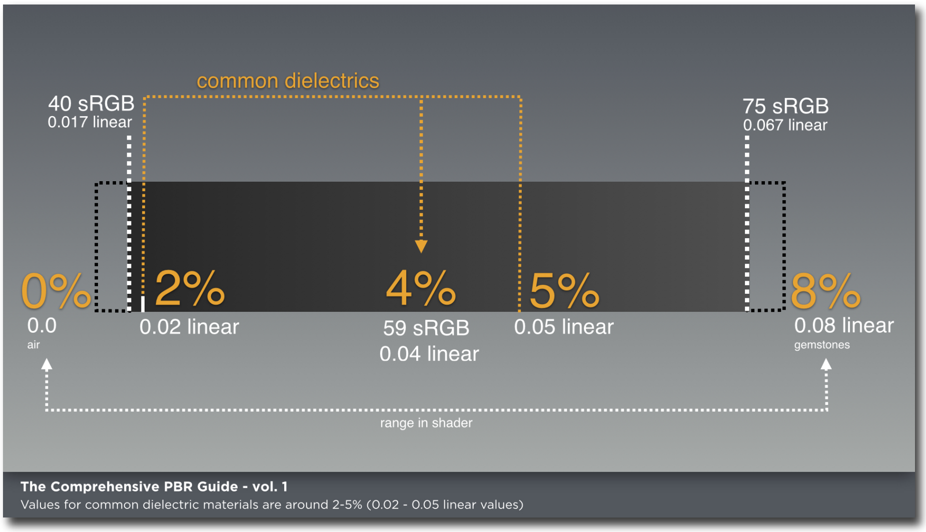


Figure 14

Just as with metals, we need to use real-world measured values, but it can be difficult to find an IOR for other materials that are not transparent. However, the value between most common dielectric materials doesn't change drastically, so we can utilize a few guidelines to follow in terms of reflectance values, which we will cover in volume two.

The value for common dielectrics is around 2-5% based on the F0 as computed by the Index of Refraction (IOR)

Linear Space Rendering

Linear space rendering can take up an entire article all on its own. So, we won't go in-depth into the specifics. However, the important takeaway is that computations are calculated in linear space.

Simply put, linear space rendering provides correct math for lighting calculations. It's about creating an environment that allows light to behave as it does in the real world. In linear space, the gamma is 1.0. However, for this to look correct to our eyes, the linear gamma needs to be shifted. Gamma-encoded space (sRGB) compensates for images that are displayed on a computer screen. The value of the image is adjusted for display.

When computing color values and performing operations on colors, all computations should be performed in linear space. A simple way to look at it is that if an image is to be displayed in the render such as base color or diffuse, then these maps need to be set as sRGB. What happens in Substance is that if the image is tagged as sRGB, it will be converted to linear for calculations and then set back to sRGB for display. However, when you store mathematical values that purely denote surface attributes in a texture such as roughness or metallic, then these maps must be set as linear.

Substance handles the conversion between linear/sRGB space for inputs automatically as well gamma-correction on the computed result in the rendered viewport. As the artist, you don't need to worry about the internal working of linear-space computations and conversions in the Substance pipeline. When using Substance materials via the Substance Integration plugin, the conversions for linear space are also handled automatically.

However, it's important to understand the process, as when Substance maps are utilized as exported bitmaps and not Substance materials, you may need to manually handle the conversions depending on the renderer you are using. You need to know that base color/diffuse maps are sRGB and the rest are linear.

When using Substance materials via the Substance Integration plugin, the conversions for linear space are also handled automatically.

Key Factors

Now that we have explored the basic theory behind the physics, we can derive some key factors for PBR.

1. Energy Conservation. A reflected ray is never brighter than the value it had when it first hit the surface. Energy Conservation is handled by the shader.

2. Fresnel. The BRDF is handled by the shader. The F0 reflectance value has minimal change for most common dielectrics and falls within a range of 2% - 5%. The F0 for metals is a high value ranging from 70-100%.

3. Specular intensity is controlled through the BRDF, roughness or glossiness map and the F0 reflectance value.

4. Lighting calculations are computed in linear space. All maps that have gamma-encoded values such as base color or diffuse are usually converted by the shader to linear, but you may have to make sure that the conversion is properly done by checking the appropriate option when importing the image in your game engine or renderer. Maps that describe surface attributes such as roughness, glossiness, metallic and height should be set to be interpreted as linear.

References

1. Physically-Based Shading at Disney Brent Burley, Walt Disney Animation Studios.
   https://disney-animation.s3.amazonaws.com/library/s2012_pbs_disney_brdf_notes_v2.pdf
2. Microfacet Models for Refraction through Rough Surfaces
   http://www.cs.cornell.edu/~srm/publications/EGSR07-btdf.pdf
3. Feeding a Physically-Based Shading Model by Sebastien Lagarde
   http://seblagarde.wordpress.com/2011/08/17/feeding-a-physical-based-lighting-mode/
4. An Introduction to BRDF Models by Daniël Jimenez Kwast
   http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Jimenez-Kwast-Daniel.pdf