Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs

Eric Heitz

INRIA ; CNRS ; Univ. Grenoble Alpes

SIGGRAPH2014

2014-09-01

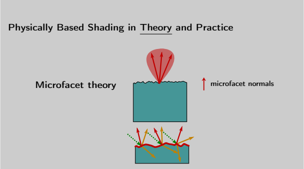
Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs

Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs

Eric Heitz

INRIA ; CNRS ; Univ. Grenoble Alpes

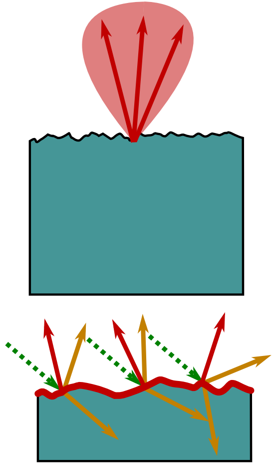
SIGGRAPH2014



Introduction

Physically Based Shading in Theory and Practice

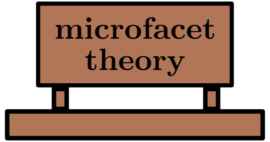
Microfacet theory



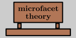
microfacet normals

In this section of the course, I will be reviewing the theoretical aspects of physically based shading, and microfacet theory in particular.

Introduction



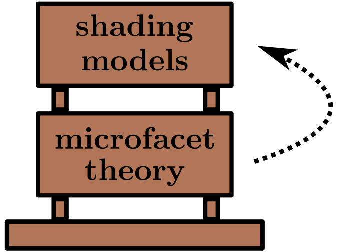
Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs

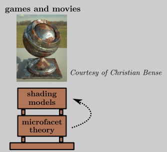


Introduction

games and movies

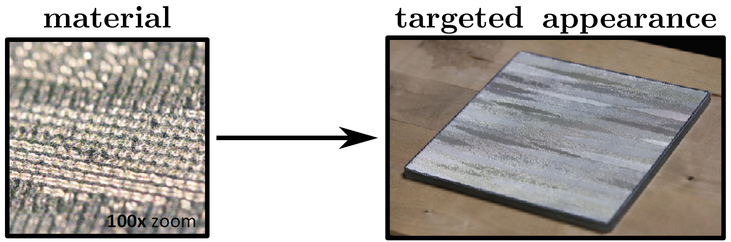
Courtesy of Christian Bense



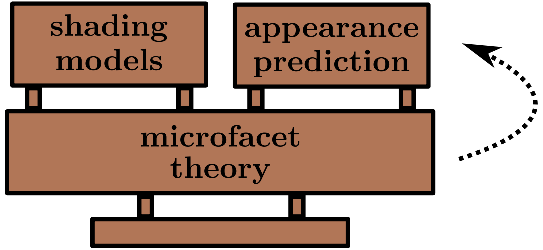


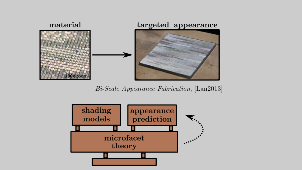
Microfacet theory is fundamental to the design of physically based shading models.

Introduction



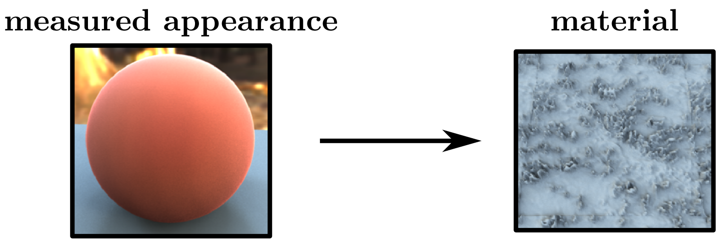
Bi-Scale Appearance Fabrication, [Lan2013]



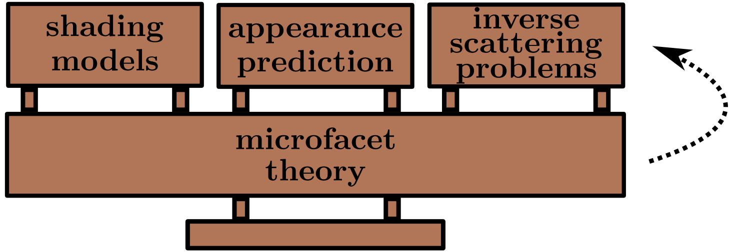


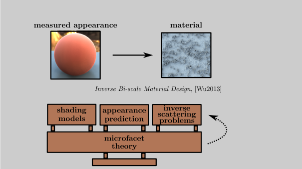
But it is not only limited to that. It is used in other applications, such as fabrication...

Introduction



Inverse Bi-scale Material Design, [Wu2013]

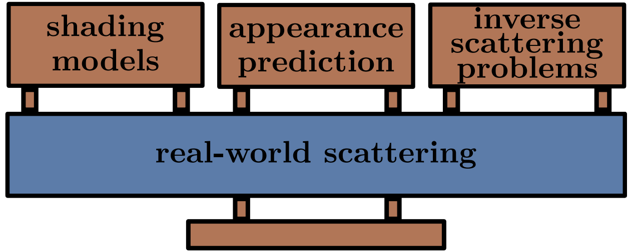


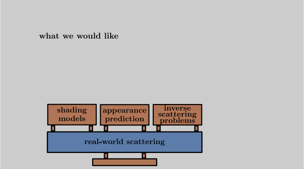


...or inverse scattering problems.

Introduction

what we would like

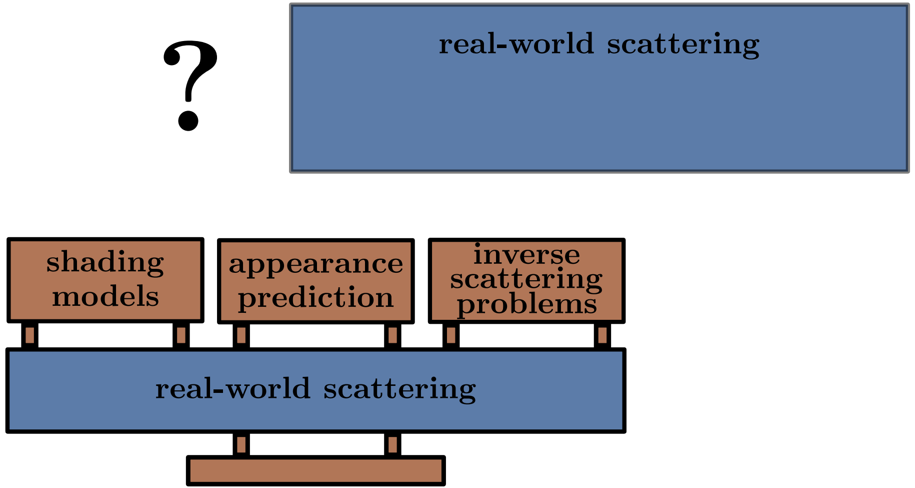


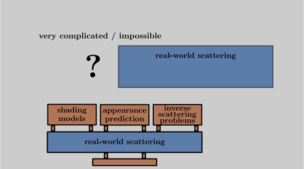


Ideally, these applications would build on a perfect descriptor of real-world scattering.

Introduction

very complicated / impossible

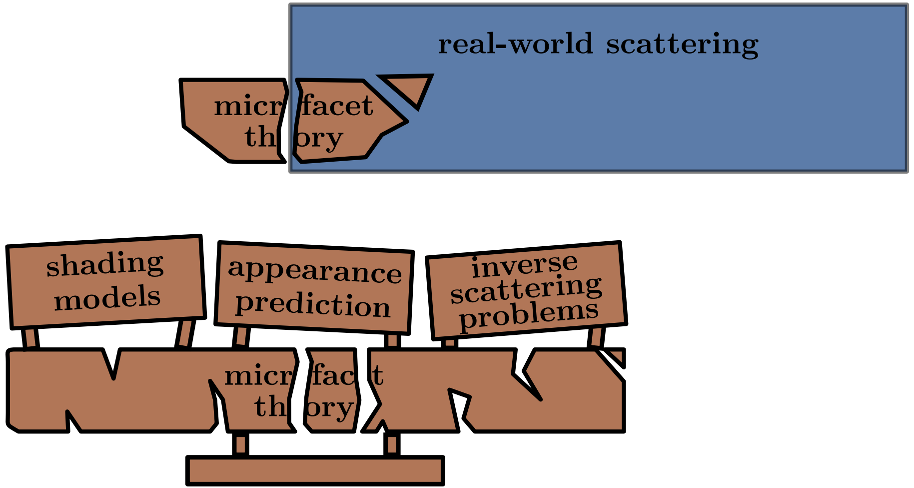


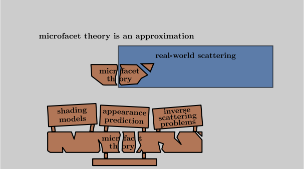


But this is either too complicated or impossible.

Introduction

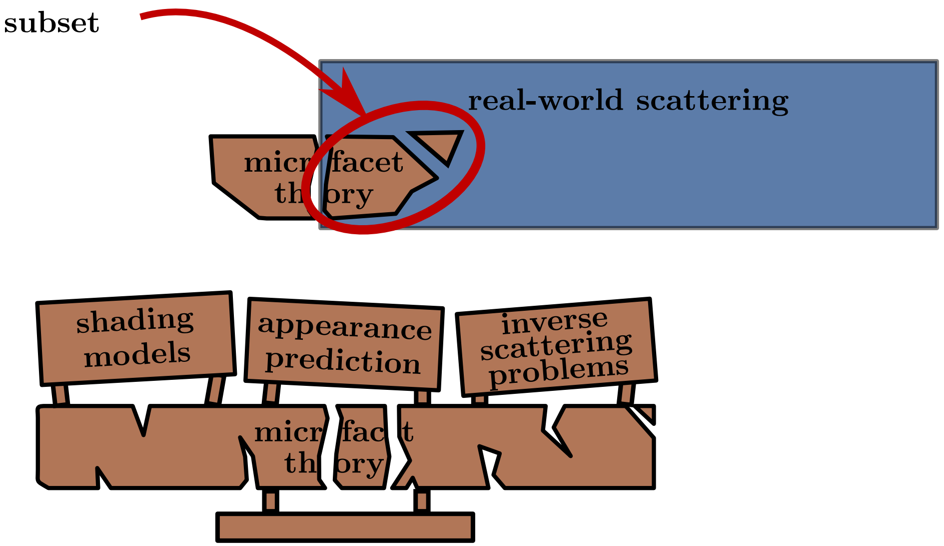
microfacet theory is an approximation

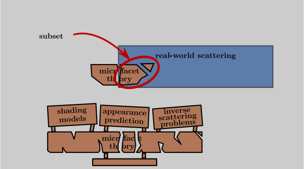




So instead we use microfacet theory as an approximation of real-world scattering. Of course, as an approximation, it comes with several limitations.

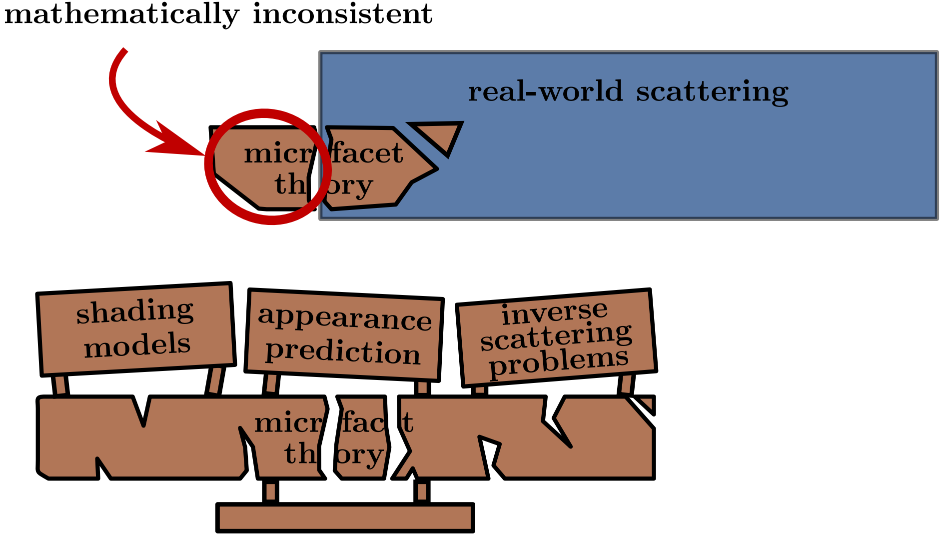
Introduction

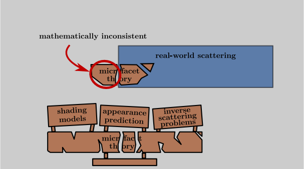




The first and most obvious limitation is that microfacet theory is, and ever will be, a subset of real-world scattering: there are some materials that cannot be described by microfacet theory. But this is reasonable, as we cannot expect a single theory to describe the entire universe.

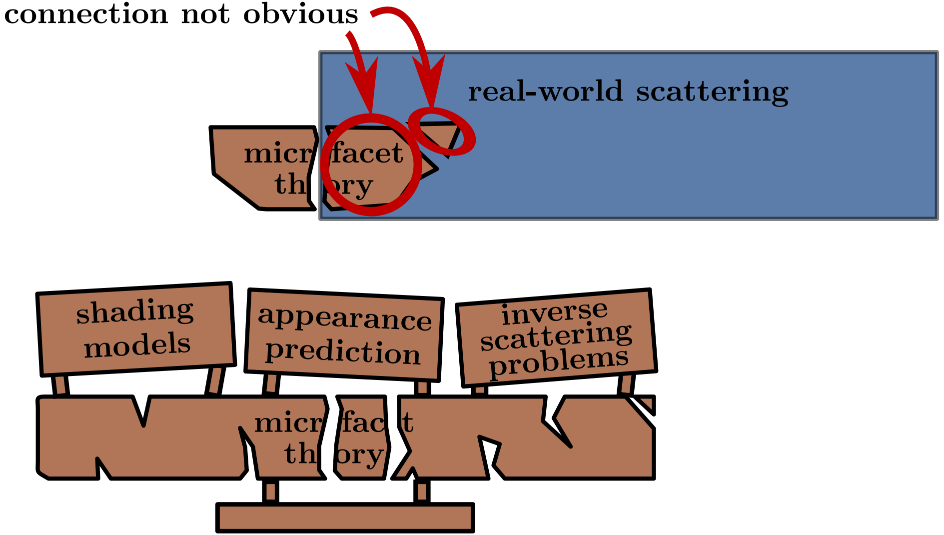
Introduction

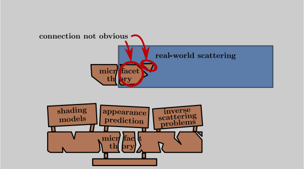




Another, more serious limitation, is that some so-called “microfacet models” are actually mathematically inconsistent.

Introduction

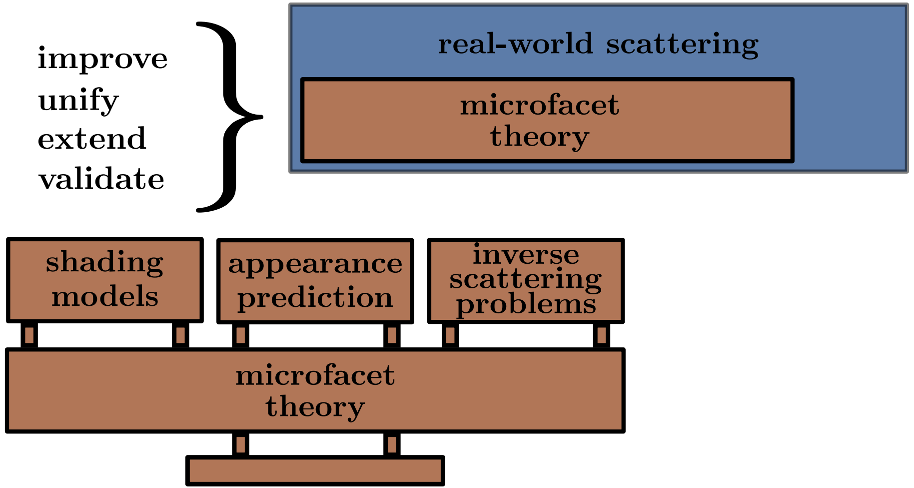


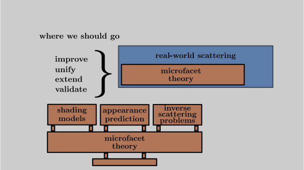


A further difficulty is that we now have so many microfacet models in the field of computer graphics that understanding how they are connected together is not always obvious.

Introduction

where we should go

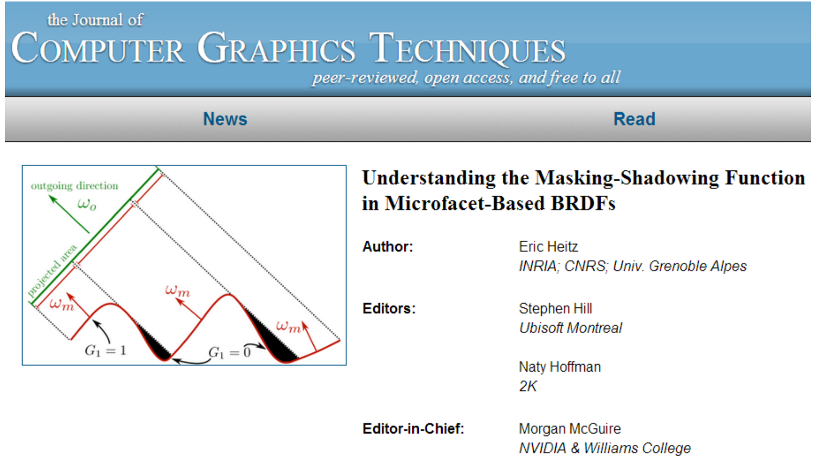




So, microfacet theory is far from perfect, but it is still one of the best tools we have at our disposal for investigating surface scattering. This is why it is important to keep studying and improving it.

Introduction

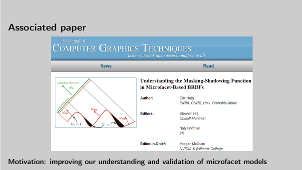
Associated paper



Motivation: improving our understanding and validation of microfacet models

2014-09-01

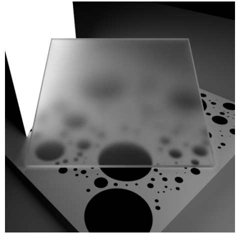
Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs



This JCGT paper serves as the course notes for this talk and we refer the reader to this paper for the derivations of the equations presented in the slides. The main motivation behind the paper is to improve how we choose and validate microfacet models. We will see that this is closely related to the choice of masking function.

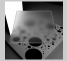
Another related paper

Importance Sampling Microfacet-Based BSDFs
using the Distribution of Visible Normals
Eric Heitz & Eugene d'Eon
EGSR 2014



→ a practical follow up of the theoretical knowledge presented in this course

Another related paper
Importance Sampling Microfacet-Based BSDFs
using the Distribution of Visible Normals
Eric Heitz & Eugene d'Eon
EGSR 2014

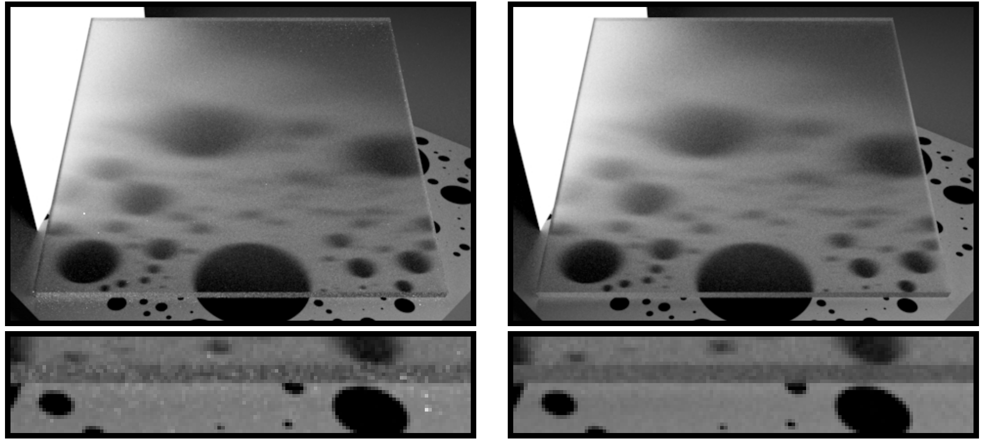


→ a practical follow up of the theoretical knowledge presented in this course

This paper will not be discussed in this talk, but is a practical follow up of its theoretical content. It is a typical example of how theoretical investigations can have practical consequences: by using our understanding of the microfacet model, we were able to design a new importance sampling technique for microfacet BSDFs.

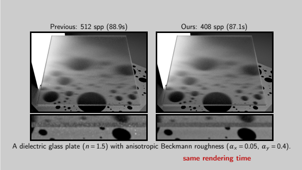
Introduction

Previous: 512 spp (88.9s) Ours: 408 spp (87.1s)



A dielectric glass plate ($n = 1.5$) with anisotropic Beckmann roughness ($\alpha_x = 0.05$, $\alpha_y = 0.4$).

same rendering time



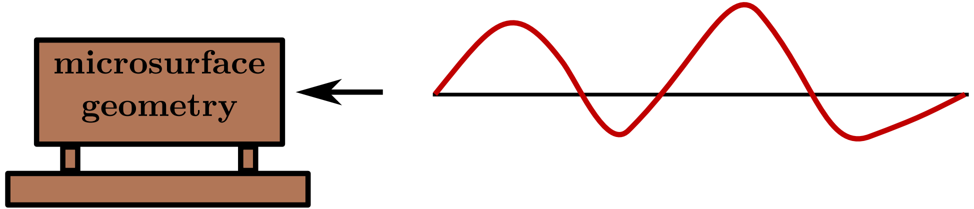
Results from the EGSR sampling paper: for the same rendering time, our technique produces images with less variance.

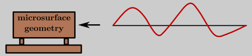
Overview of Microfacet Theory and Related Problems

Overview of Microfacet Theory and Related Problems

To get started, I will give you an overview of how microfacet theory is constructed.

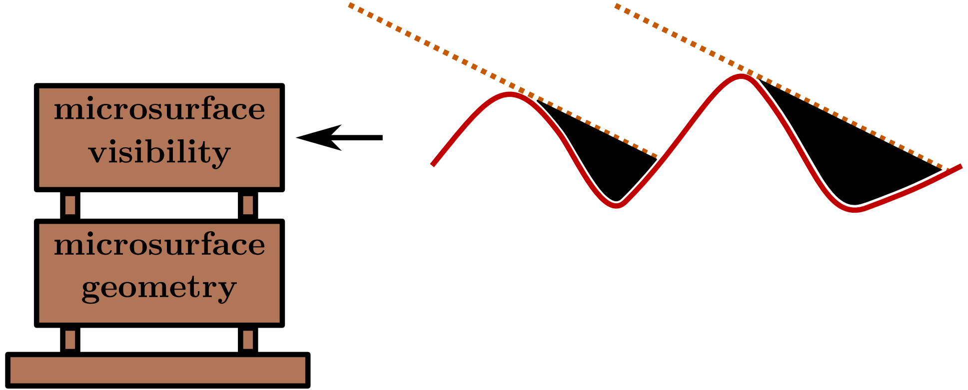
Overview of Microfacet Theory and Related Problems

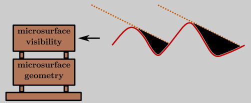




While the geometric surface (or macrosurface) may appear flat, microfacet theory assumes that a very small and rough microsurface is responsible for the scattering occurring at the material interface. The first step is to model the geometry of this microsurface, in other words: what does it look like?

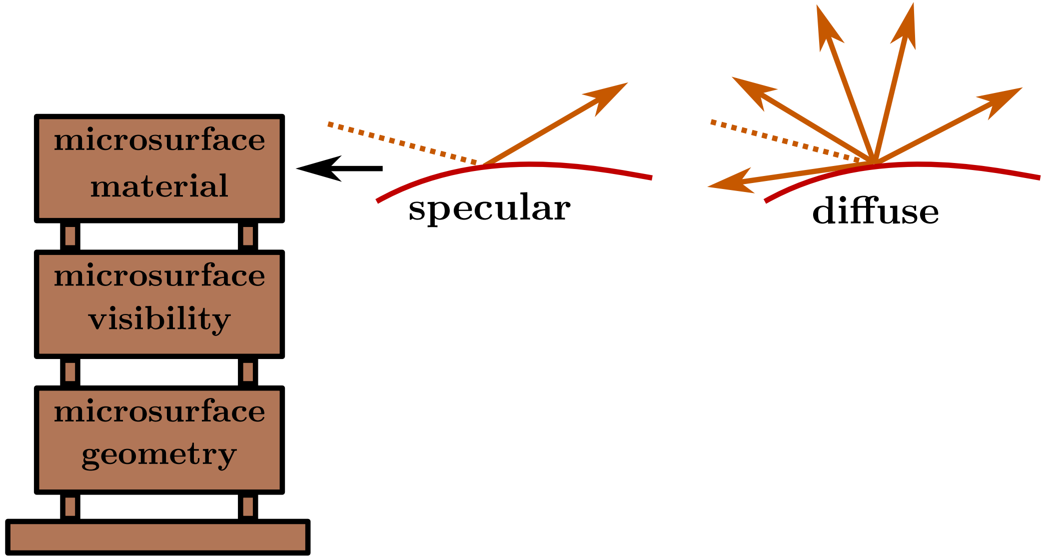
Overview of Microfacet Theory and Related Problems

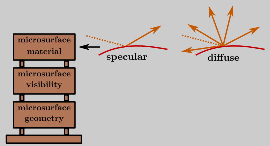




Once the geometry of the microsurface has been established, we can compute what parts of it will be visible for a given view direction.

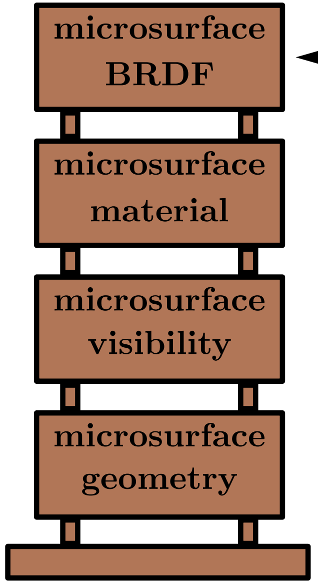
Overview of Microfacet Theory and Related Problems

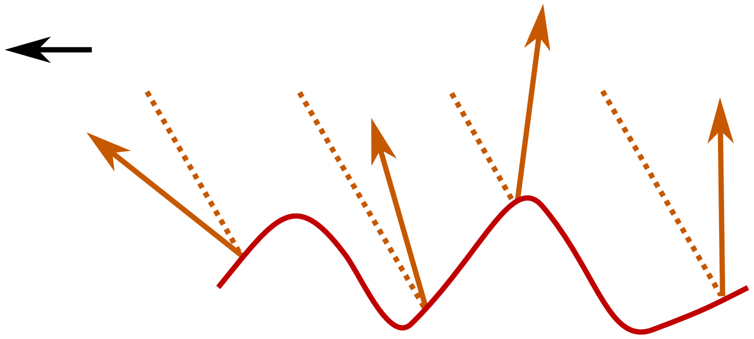




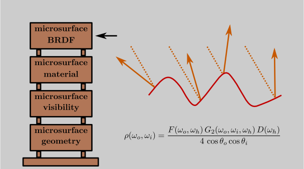
Once we know what parts of the microsurface are visible, we need to model how they will be interacting with the light. Usually, microfacet models assume that the microfacets are perfectly specular and produce mirror-like reflections. Other models, like Oren and Nayar's, assume that the microfacets are perfect diffusers.

Overview of Microfacet Theory and Related Problems

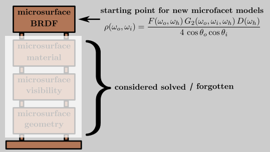




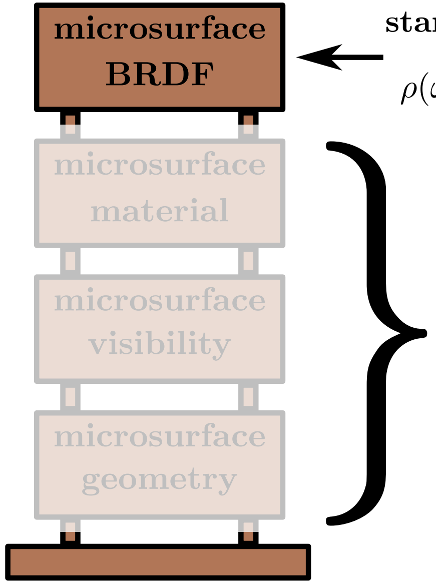
$$\rho(\omega_o, \omega_i) = \frac{F(\omega_o, \omega_h)\, G_2(\omega_o, \omega_i, \omega_h)\, D(\omega_h)}{4\, \cos\theta_o \cos\theta_i}$$



Once the microsurface geometry, visibility and material are fixed, we can finally derive the complete microsurface BRDF expression. In the case of specular microfacets, this leads to the famous Cook and Torrance equation.



Overview of Microfacet Theory and Related Problems



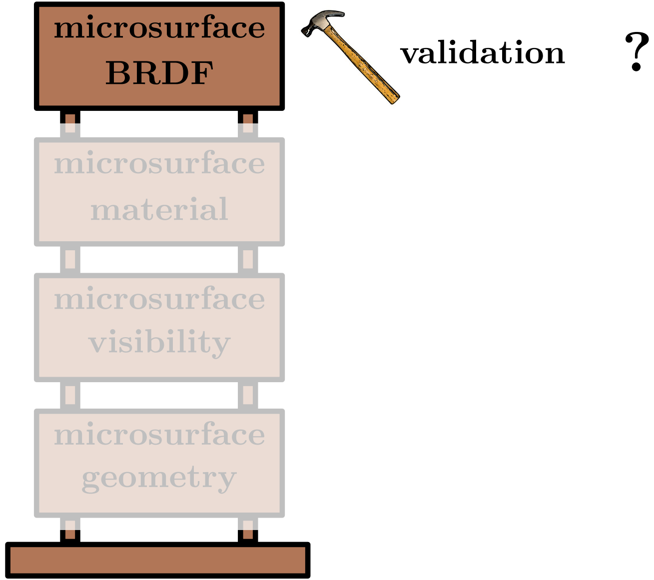
starting point for new microfacet models

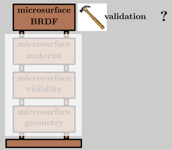
$$\rho(\omega_o, \omega_i) = \frac{F(\omega_o, \omega_h)\, G_2(\omega_o, \omega_i, \omega_h)\, D(\omega_h)}{4 \cos\theta_o \cos\theta_i}$$

considered solved / forgotten

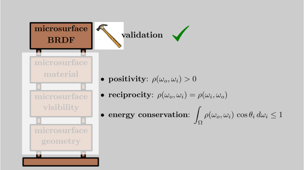
In common microfacet papers, the first three derivation steps are considered "previous work", and the Cook and Torrance equation usually serves as the starting point from which to derive new models. By modifying F, G_2 and D, it is possible to create a wide variety of different models.

Overview of Microfacet Theory and Related Problems

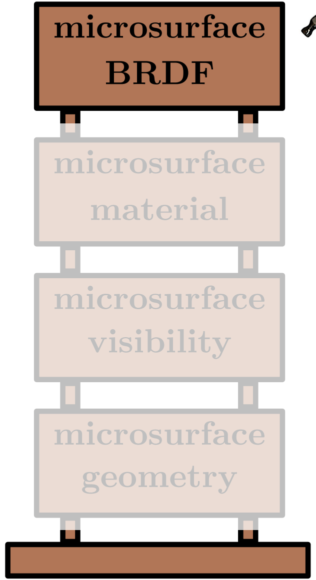


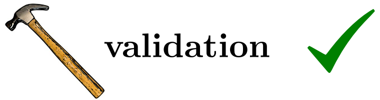


Once a new model has been created, it has to be validated.



Overview of Microfacet Theory and Related Problems

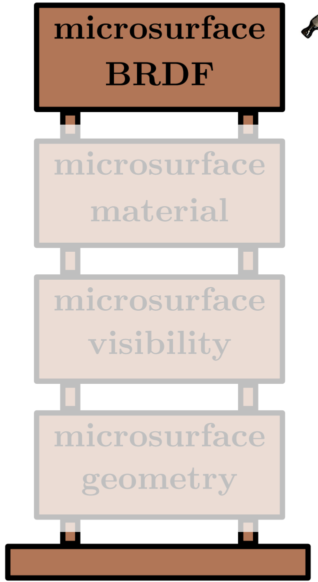




- **positivity**: $\rho(\omega_o, \omega_i) > 0$
- **reciprocity**: $\rho(\omega_o, \omega_i) = \rho(\omega_i, \omega_o)$
- **energy conservation**: $\int_\Omega \rho(\omega_o, \omega_i)\, \cos\theta_i \, d\omega_i \leq 1$

We usually consider a microfacet model to be "physically based" if it is positive, reciprocal, and energy conserving.

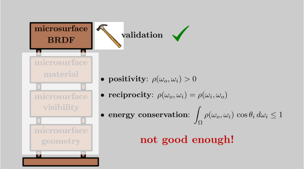
Overview of Microfacet Theory and Related Problems



validation ✓

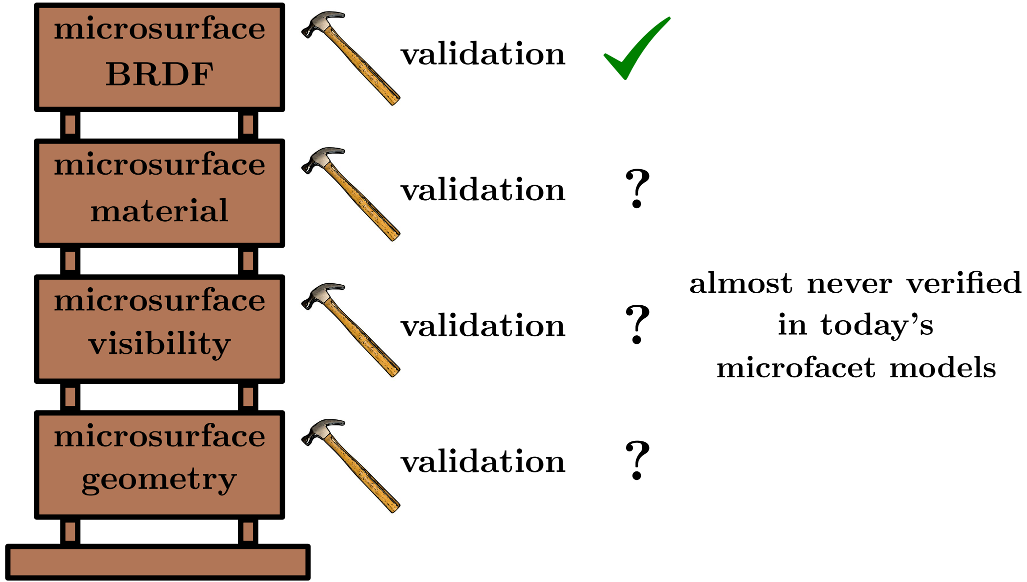
- **positivity**: $\rho(\omega_o, \omega_i) > 0$
- **reciprocity**: $\rho(\omega_o, \omega_i) = \rho(\omega_i, \omega_o)$
- **energy conservation**: $\int_\Omega \rho(\omega_o, \omega_i) \cos\theta_i \, d\omega_i \leq 1$

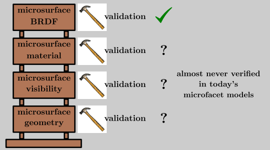
not good enough!



However, those criteria are not sufficient to validate a new model, because they are not restrictive enough. Intuitively, one could come up with some random BRDF model that easily satisfies those three conditions, and yet fails to relate to any meaningful physical model.

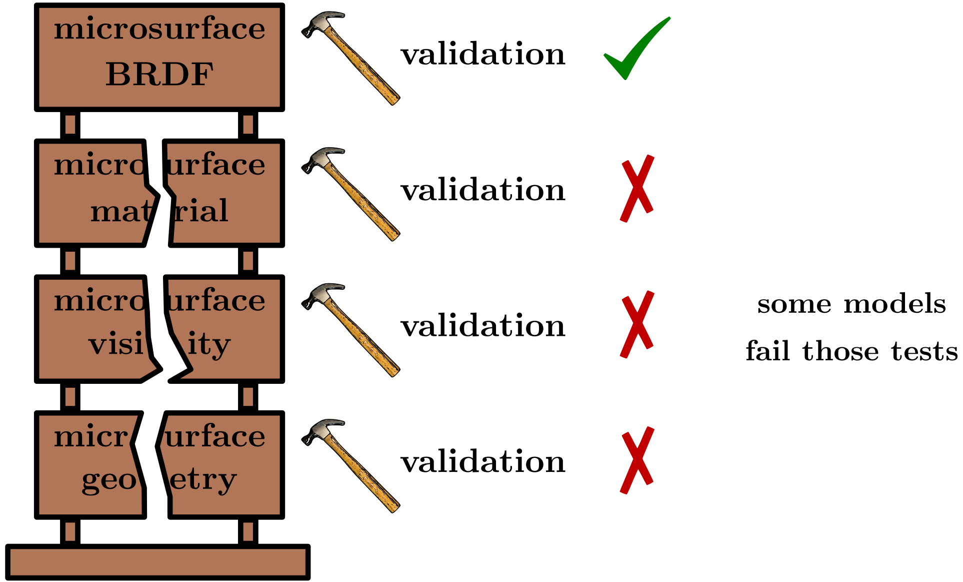
Overview of Microfacet Theory and Related Problems

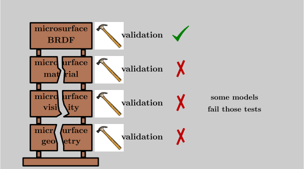




What's missing? The problem is that these intermediate derivation steps have also their own validation criteria. However, since they are almost never mentioned, these associated criteria are almost never checked.

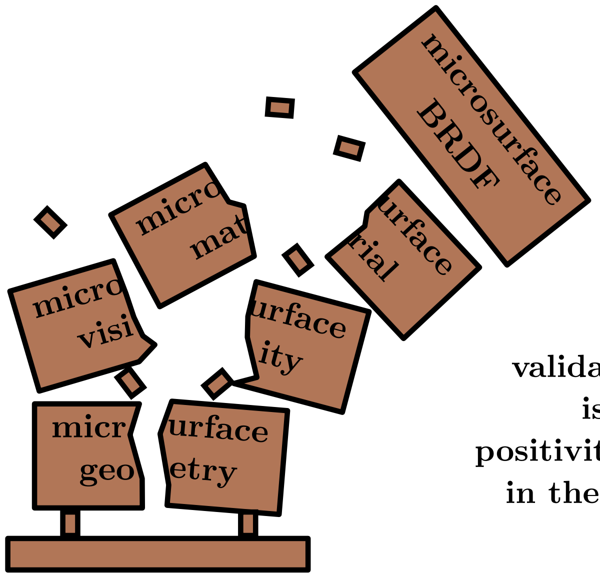
Overview of Microfacet Theory and Related Problems





It turns out that, within the set of what we call "microfacet models" today, there are some that don't validate those criteria. Such models should not be called "microfacet based", nor "physically based".

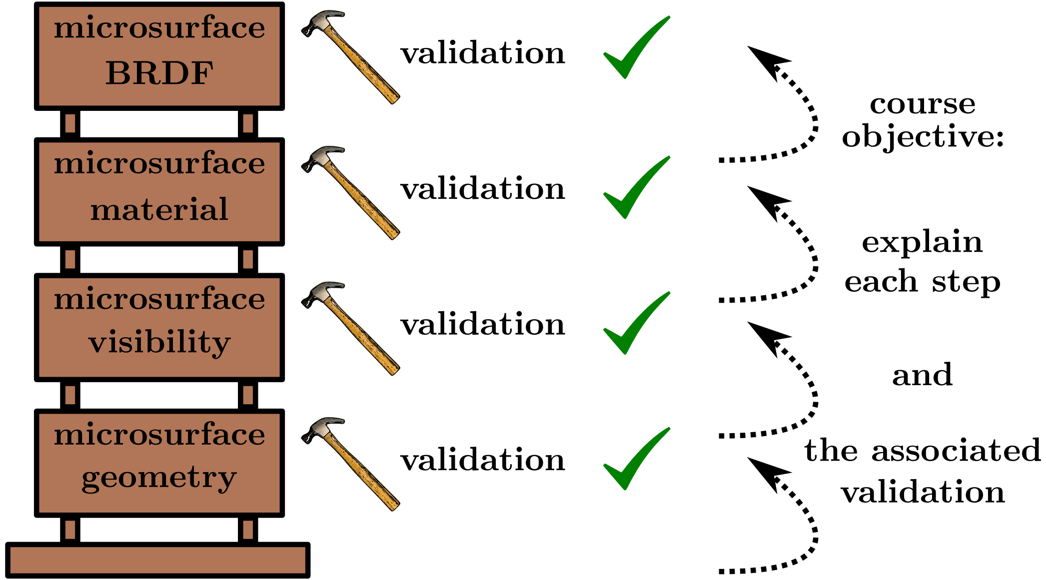
Overview of Microfacet Theory and Related Problems

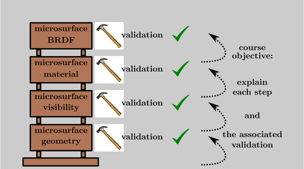


validating microfacet models
is not only checking
positivity, reciprocity and energy
in the final BRDF expression

This is probably the main message of this talk.

Overview of Microfacet Theory and Related Problems





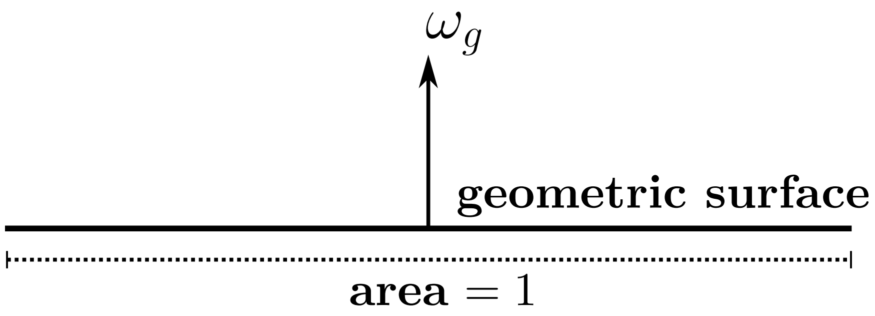
The objective of this presentation (and the course notes) is to review those derivations and for each one determine the associated validation criteria. Then, we will use them to assess common models.

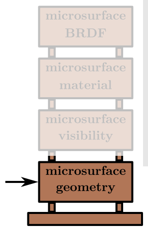
The microfacet model

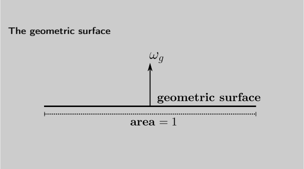
The microfacet model

The microfacet model

The geometric surface



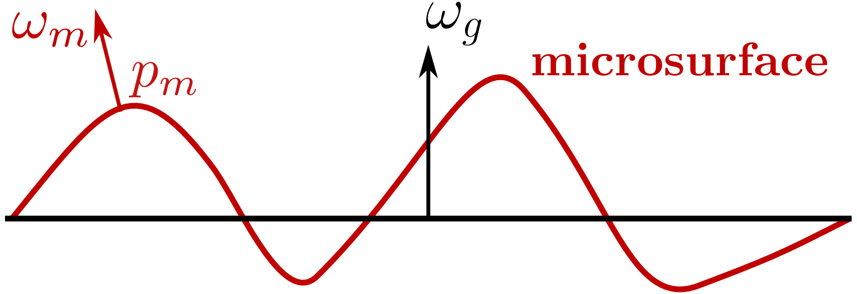


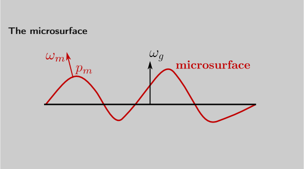


Microfacet theory starts with the geometric surface. In the case of a triangle mesh that we wish to shade, the "geometric surface" refers to a very small and locally planar piece of this mesh. Its area is 1 by convention.

The microfacet model

The microsurface

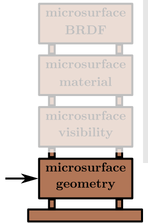




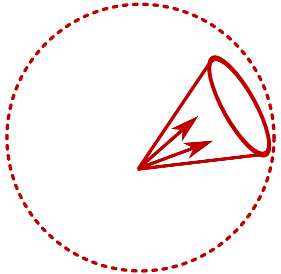
Next, we assume that what is actually interacting with the light is not the geometric surface, but a rough microsurface, composed of microfacets. At this point, the scattering occurring at the object interface can be described as a spatial function defined on the microsurface.

The microfacet model

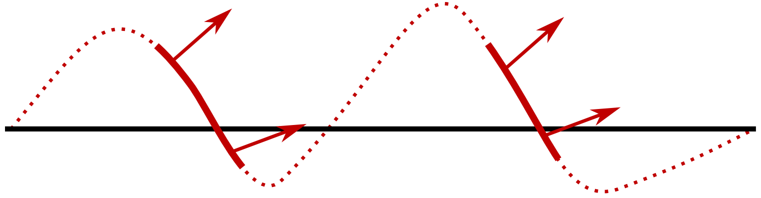
solid angle (sr)



The distribution of normals $D(\omega_m)$



solid angle (sr)



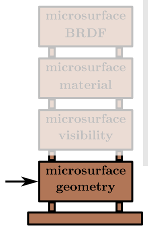
area (m^2)

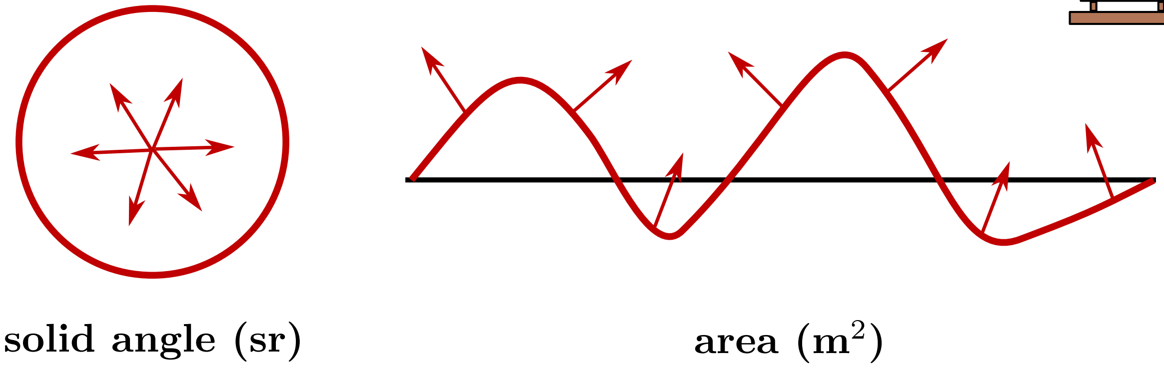
area (m^2)

However, working with a spatial description of the problem is needlessly complicated. It is much easier to use a statistical description that's defined on the sphere. This is what the distribution of normals is for: it relates a spatial measure defined on the microsurface to a statistical measure defined on the sphere.

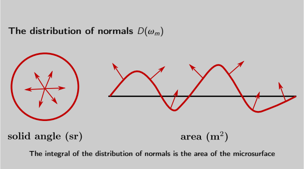
⚠This slide is animated (works with Acrobat Reader).

The distribution of normals $D(\omega_m)$





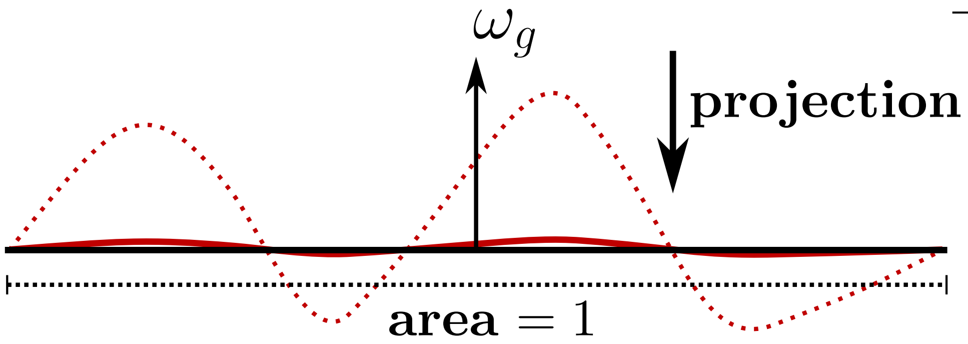
The integral of the distribution of normals is the area of the microsurface



As a consequence, the measure of the entire distribution of normals (its integral) is the measure of the entire microsurface (its area).

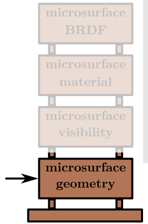
The microfacet model

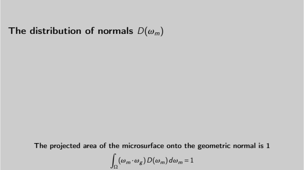
The distribution of normals $D(\omega_m)$



The projected area of the microsurface onto the geometric normal is 1

$$\int_{\Omega} (\omega_m \cdot \omega_g)\, D(\omega_m)\, d\omega_m = 1$$

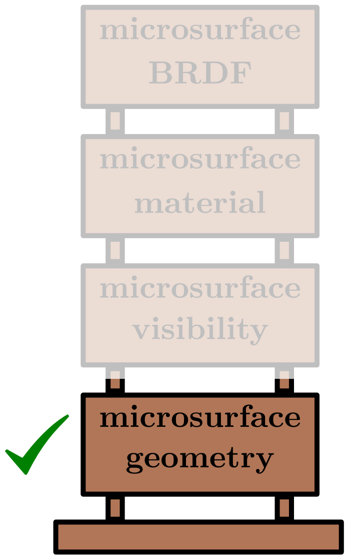




Since the distribution of normals is a statistical descriptor of the microsurface, it should precisely obey the same properties. The first property is the conservation of the projected area: the microsurface projected onto the geometric surface is the geometric surface, and so the projected area of the microsurface is the area of the geometric surface (1 by convention, as mentioned earlier).

⚠This slide is animated (works with Acrobat Reader).

Validation equations



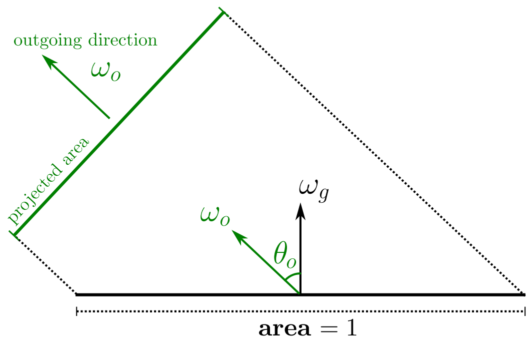
$$\int_\Omega (\omega_m \cdot \omega_g)\, D(\omega_m)\, d\omega_m = 1$$

$\int_\Omega (\omega_m \cdot \omega_g) D(\omega_m) d\omega_m = 1$

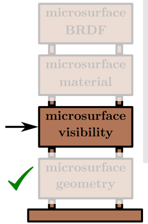
We can use this property for validation.

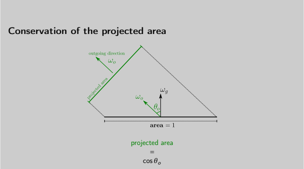
The microfacet model

Conservation of the projected area



projected area
=
$\cos\theta_o$

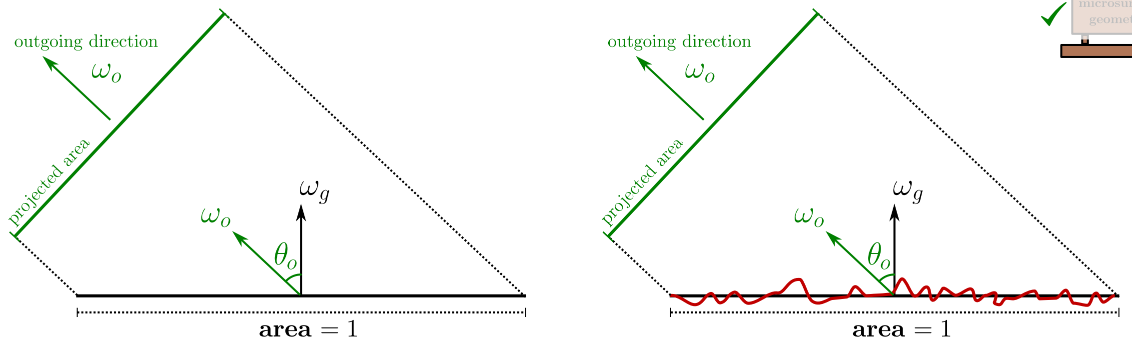




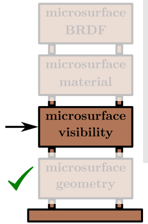
We can generalize the concept of the projected area to any direction. The projected area of the geometric surface is the cosine of the projection direction.

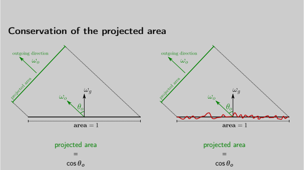
The microfacet model

Conservation of the projected area



projected area = $\cos\theta_o$	projected area = $\cos\theta_o$

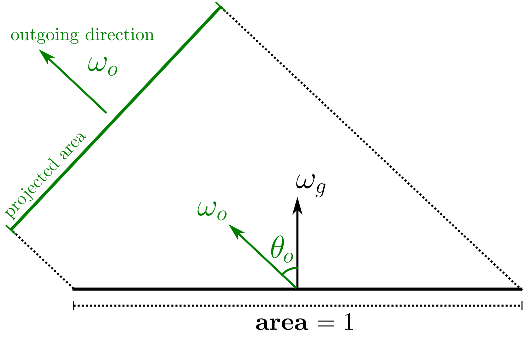




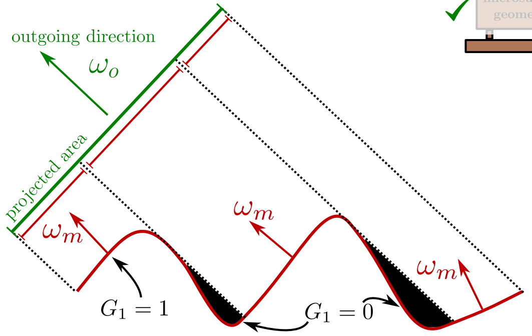
If we replace the geometric surface by the microsurface in this figure, it may appear that the projected area doesn't change. However, this is only because the microsurface features are small.

The microfacet model

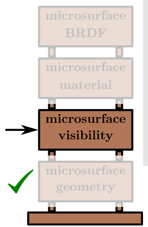
Conservation of the projected area

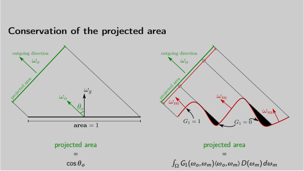


projected area
=
$\cos\theta_o$



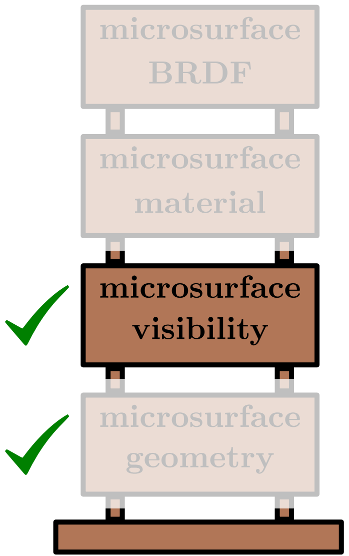
projected area
=
$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m$





By zooming into the microsurface, we can see that its projected area is the sum of the projected areas of the microfacets that are visible. At this point, we need to introduce a visbility term, G_1, to discard the microfacets that are occluded. This is the masking function. Thanks to this masking function, we can write a new equation.

Validation equations



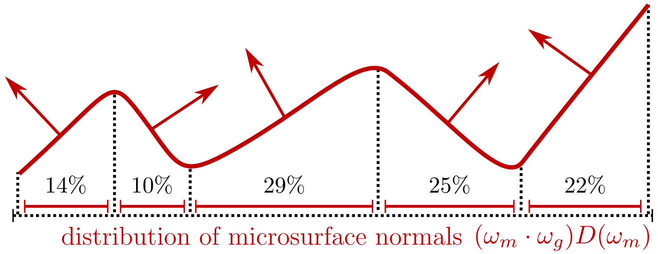
$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

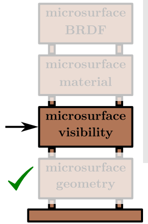
$$\int_\Omega (\omega_m\cdot\omega_g)\, D(\omega_m)\, d\omega_m = 1$$

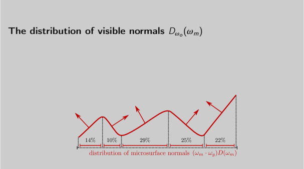
We can add this equation to our validation list.

The microfacet model

The distribution of visible normals $D_{\omega_o}(\omega_m)$

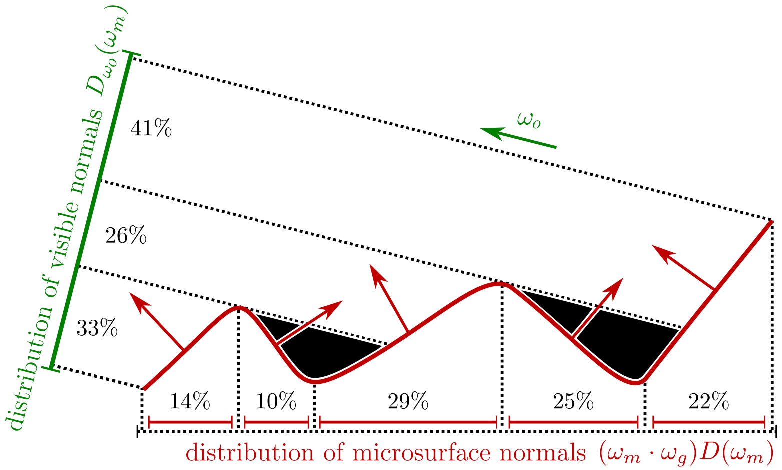




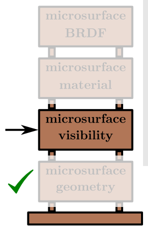


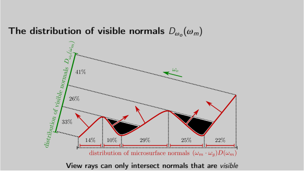
As we have seen, the microsurface is described by the distribution of normals.

The distribution of visible normals $D_{\omega_o}(\omega_m)$



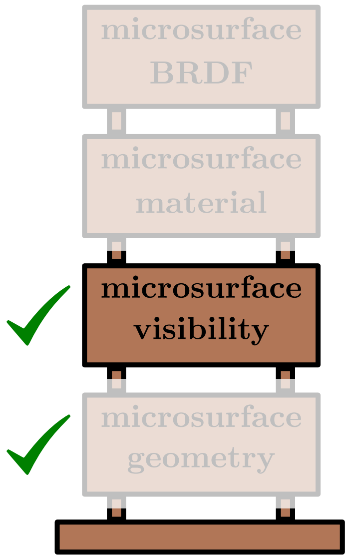
View rays can only intersect normals that are ***visible***





However, only the microfacets that are visible will reflect light towards the viewer. Incorporating visibility into the distribution gives us the <u>distribution of visible normals</u>. If the model is well designed, this distribution should be normalized, in the sense that the percentages shown in the figure should add up to exactly 1.

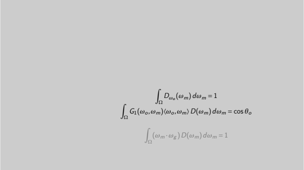
Validation equations



$$\int_\Omega D_{\omega_o}(\omega_m)\, d\omega_m = 1$$

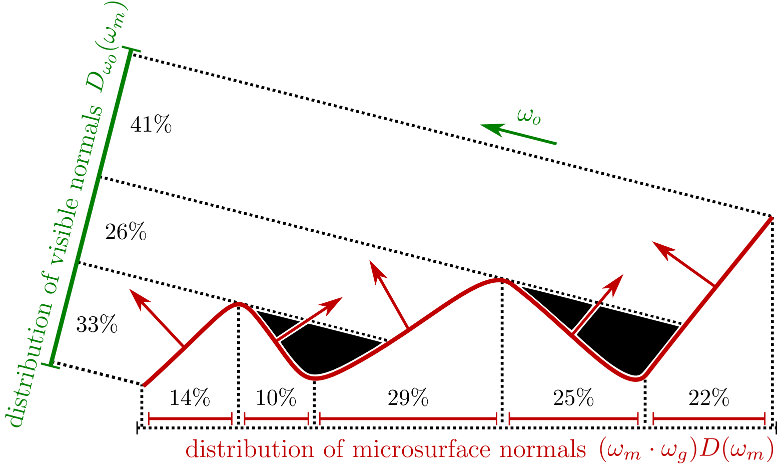
$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

$$\int_\Omega (\omega_m\cdot\omega_g)\, D(\omega_m)\, d\omega_m = 1$$

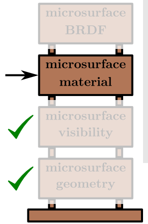


The microfacet model

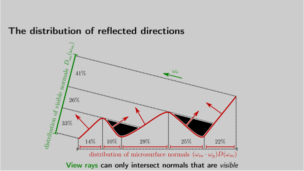
The distribution of reflected directions



View rays can only intersect normals that are *visible*



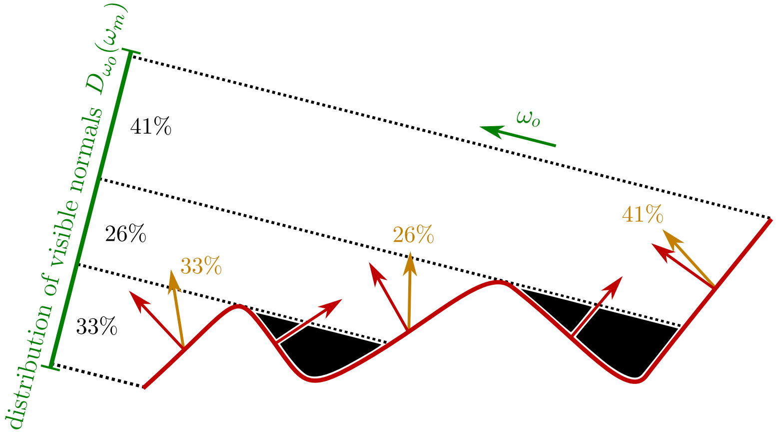
2014-09-01



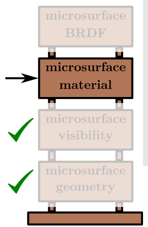
If we return to this figure that shows the microfacets that are visible...

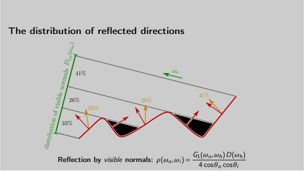
The microfacet model

The distribution of reflected directions



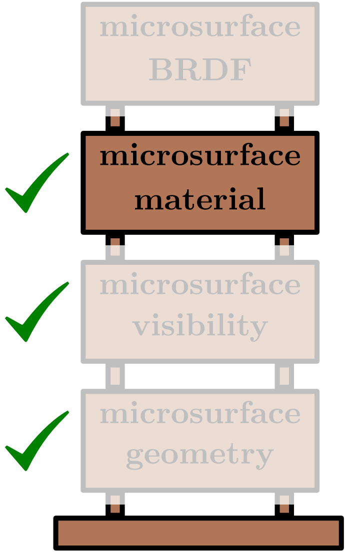
Reflection by ***visible*** **normals:** $\rho(\omega_o,\omega_i) = \dfrac{G_1(\omega_o,\omega_h)\,D(\omega_h)}{4\cos\theta_o\cos\theta_i}$





...and apply a light transport operator such as specular reflection, we get the equation of an incomplete BRDF model. At this point in the model, no energy is lost, and we can see that the distribution of reflected directions (the percentages in orange) exactly matches the distribution of visible microfacets (the percentages in black). Hence, this incomplete BRDF model should be normalized as well.

Validation equations

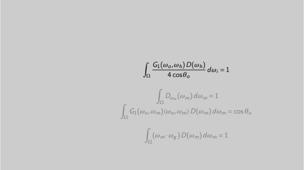


$$\int_\Omega \frac{G_1(\omega_o,\omega_h)\, D(\omega_h)}{4\cos\theta_o}\, d\omega_i = 1$$

$$\int_\Omega D_{\omega_o}(\omega_m)\, d\omega_m = 1$$

$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

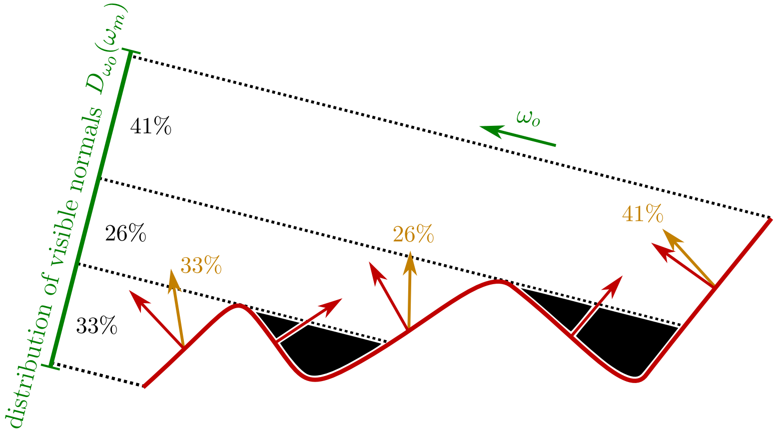
$$\int_\Omega (\omega_m\cdot\omega_g)\, D(\omega_m)\, d\omega_m = 1$$



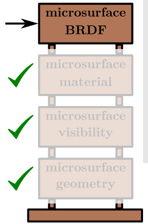
We have established validation criteria for these three intermediate steps. We can now derive the entire BRDF expression by adding the missing terms.

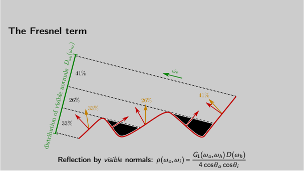
The microfacet model

The Fresnel term



Reflection by *visible* normals: $\rho(\omega_o,\omega_i) = \dfrac{G_1(\omega_o,\omega_h)\, D(\omega_h)}{4\cos\theta_o \cos\theta_i}$

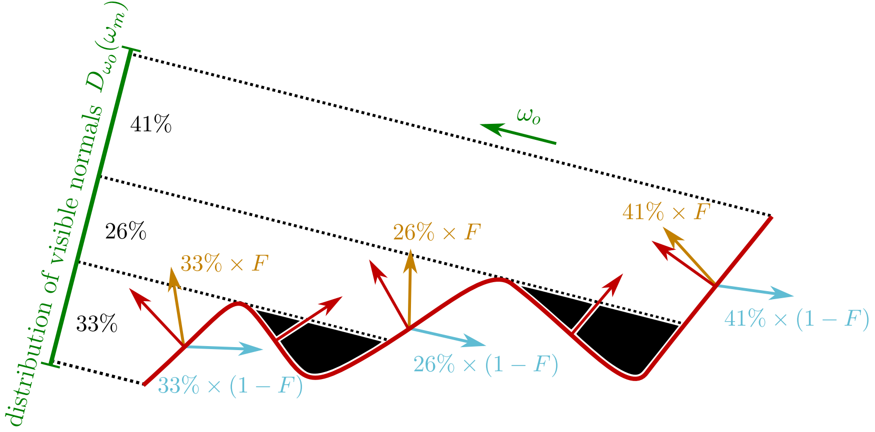




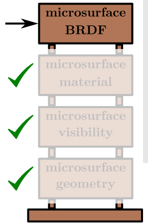
In this figure we assumed a perfect specular reflection.

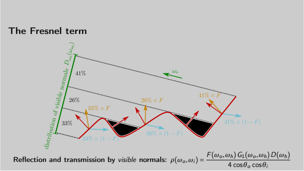
The microfacet model

The Fresnel term

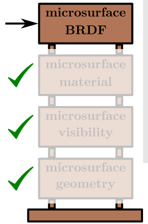


Reflection and transmission by *visible* normals: $\rho(\omega_o,\omega_i) = \dfrac{F(\omega_o,\omega_h)\, G_1(\omega_o,\omega_h)\, D(\omega_h)}{4\cos\theta_o \cos\theta_i}$

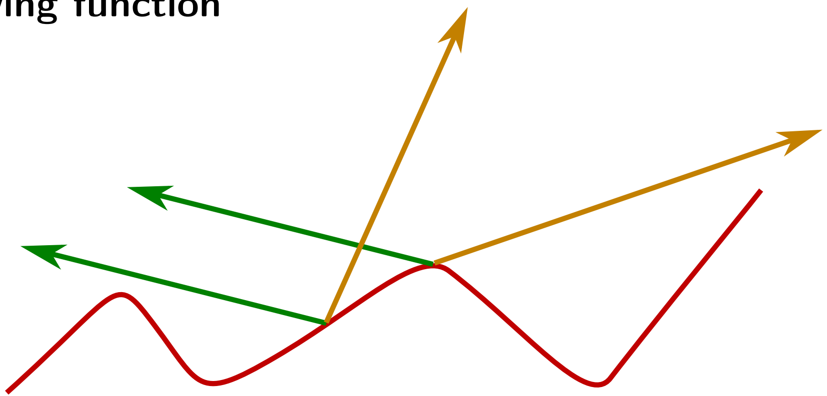




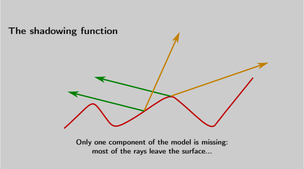
However, on a physical surface, only some of the energy is reflected. The rest is either transmitted or absorbed. This is modeled by introducing the Fresnel term F into the equation.



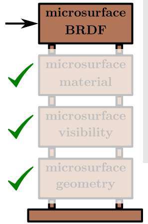
The shadowing function



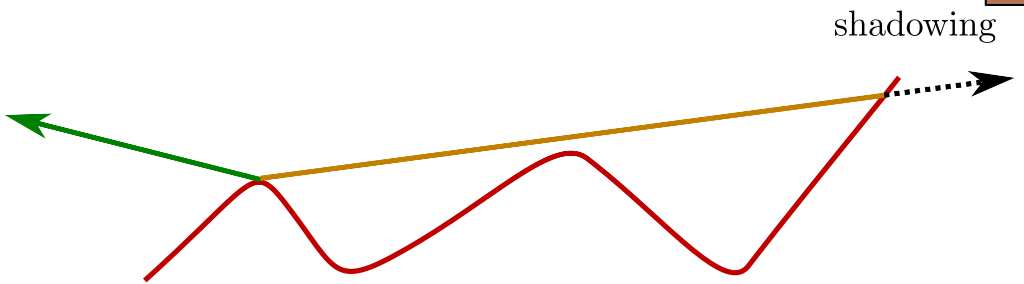
Only one component of the model is missing:
most of the rays leave the surface...



We are now very close to the final model. However, one compenent is still missing, and this is because, after the reflection some of the rays will leave the surface...

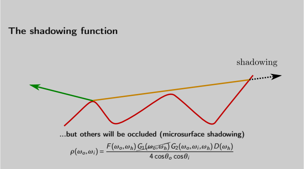


The shadowing function



...but others will be occluded (microsurface shadowing)

$$\rho(\omega_o,\omega_i) = \frac{F(\omega_o,\omega_h)\,\cancel{G_1(\omega_o,\omega_h)}\,G_2(\omega_o,\omega_i,\omega_h)\,D(\omega_h)}{4\cos\theta_o\cos\theta_i}$$



...but others will be occluded, i.e., they will intersect the surface a second time. This is called microsurface shadowing. To incorporate this effect into the model, we replace the masking function G_1, which represents visibility for the outgoing direction only, by a masking-shadowing function, G_2, which represents simultaneous visibility for the outgoing and incident directions.

The microfacet model

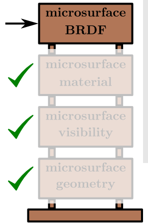
The complete microfacet BRDF model

$$\rho(\omega_o,\omega_i) = \frac{F(\omega_o,\omega_h)\, G_2(\omega_o,\omega_i,\omega_h)\, D(\omega_h)}{4\cos\theta_o \cos\theta_i}$$

Validation?

- **Positivity** $\rho(\omega_o,\omega_i) > 0$
- **Reciprocity** $\rho(\omega_o,\omega_i) = \rho(\omega_i,\omega_o)$
- **Energy conservation** $\int_\Omega \rho(\omega_o,\omega_i) \cos\theta_i \, d\omega_i \le 1$

Very weak conditions. Inappropriate for validation!



The complete microfacet BRDF model

$$\rho(\omega_o,\omega_i) = \frac{F(\omega_o,\omega_h)\, G_2(\omega_o,\omega_i,\omega_h)\, D(\omega_h)}{4\cos\theta_o \cos\theta_i}$$

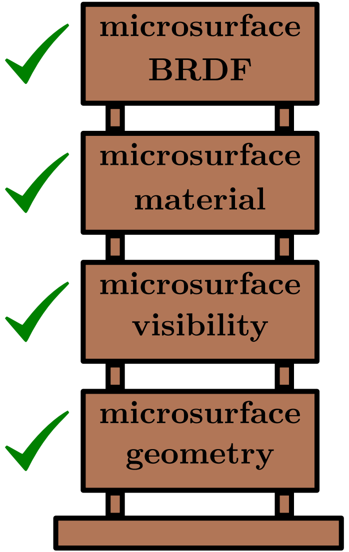
Validation?

- Positivity $\rho(\omega_o,\omega_i) > 0$
- Reciprocity $\rho(\omega_o,\omega_i) = \rho(\omega_i,\omega_o)$
- Energy conservation $\int_\Omega \rho(\omega_o,\omega_i) \cos\theta_i \, d\omega_i \le 1$

Very weak conditions. Inappropriate for validation!

We have now derived the complete BRDF model. Since the last terms we introduced – Fresnel and shadowing – are responsible for energy loss, the strict validation equations we had at the beginning have disappeared, and what is left is only a weak inequality, which is not very appropriate for thorough validation.

Validation equations



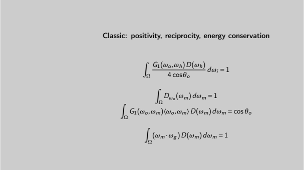
Classic: positivity, reciprocity, energy conservation

$$\int_\Omega \frac{G_1(\omega_o, \omega_h)\, D(\omega_h)}{4 \cos\theta_o}\, d\omega_i = 1$$

$$\int_\Omega D_{\omega_o}(\omega_m)\, d\omega_m = 1$$

$$\int_\Omega G_1(\omega_o, \omega_m) \langle \omega_o, \omega_m \rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

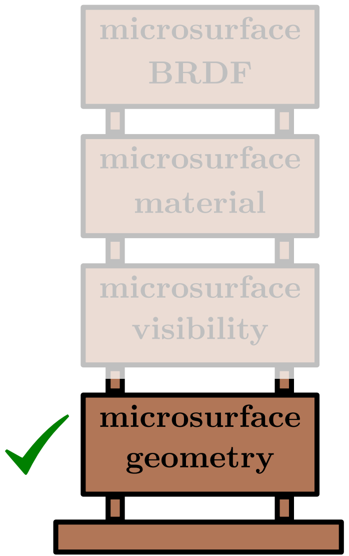
$$\int_\Omega (\omega_m \cdot \omega_g)\, D(\omega_m)\, d\omega_m = 1$$



However, on the way towards the final model, we have gathered several useful equations. We will now use them to assess common models.

Review of Common Distributions of Normals

Validation equations



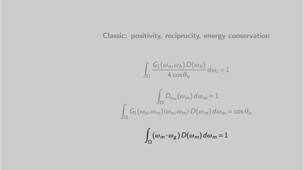
Classic: positivity, reciprocity, energy conservation

$$\int_\Omega \frac{G_1(\omega_o,\omega_h)\, D(\omega_h)}{4\cos\theta_o}\, d\omega_i = 1$$

$$\int_\Omega D_{\omega_o}(\omega_m)\, d\omega_m = 1$$

$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

$$\int_\Omega (\omega_m \cdot \omega_g)\, D(\omega_m)\, d\omega_m = 1$$

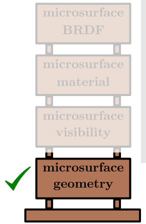


This equation can be used to validate distributions of normals.

Review of Common Distributions of Normals

Common distributions of normals

Name	Validation?	“physically based”?
Blinn-Phong	✗	✗
Normalized Blinn-Phong	✓	✓
Ward	✗	✗
Beckmann	✓	✓
GGX	✓	✓

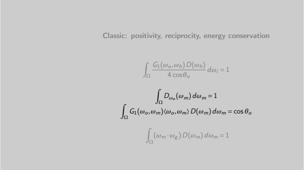


Common distributions of normals

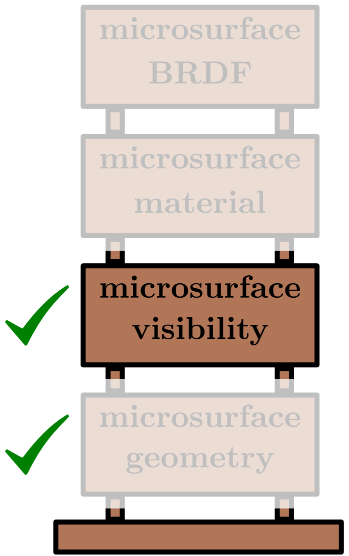
Name	Validation?	“physically based”?
Blinn-Phong	✗	✗
Normalized Blinn-Phong	✓	✓
Ward	✗	✗
Beckmann	✓	✓
GGX	✓	✓

Here are some examples (not exhaustive) of common distributions of normals. We can see, for instance, that the old Blinn-Phong and Ward distributions are not appropriately normalized.

Review of Common Masking Functions



Validation equations



Classic: positivity, reciprocity, energy conservation

$$\int_\Omega \frac{G_1(\omega_o,\omega_h)\, D(\omega_h)}{4\cos\theta_o}\, d\omega_i = 1$$

$$\int_\Omega D_{\omega_o}(\omega_m)\, d\omega_m = 1$$

$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

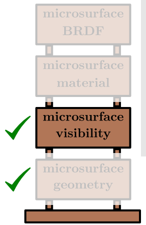
$$\int_\Omega (\omega_m \cdot \omega_g)\, D(\omega_m)\, d\omega_m = 1$$

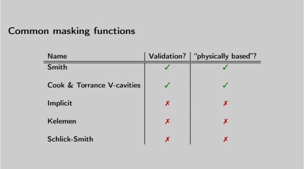
We can use these two equations to validate masking functions.

Review of Common Masking Functions

Common masking functions

Name	Validation?	“physically based”?
Smith	✓	✓
Cook & Torrance V-cavities	✓	✓
Implicit	✗	✗
Kelemen	✗	✗
Schlick-Smith	✗	✗



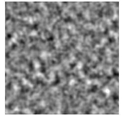
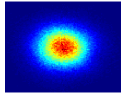
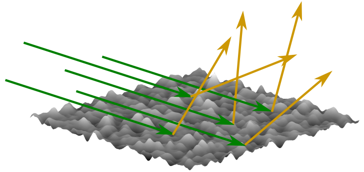
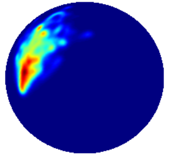

Common masking functions

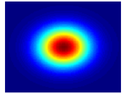
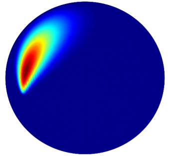
Name	Validation?	“physically based”?
Smith	✓	✓
Cook & Torrance V-cavities	✓	✓
Implicit	✗	✗
Kelemen	✗	✗
Schlick-Smith	✗	✗

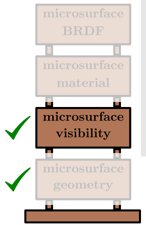
There are a lot of different masking functions in the literature. To my knowledge, only two of them satisfy the validation equations: the Smith and V-cavity masking functions. This is because they are both based on a microsurface model. The others should not be called “physically based”, in the sense that there is no possible microsurface model from which they can be derived.

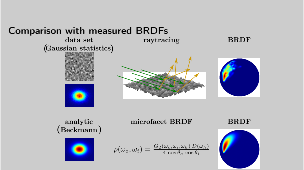
Review of Common Masking Functions

Comparison with measured BRDFs

data set (Gaussian statistics)	raytracing	BRDF
 		

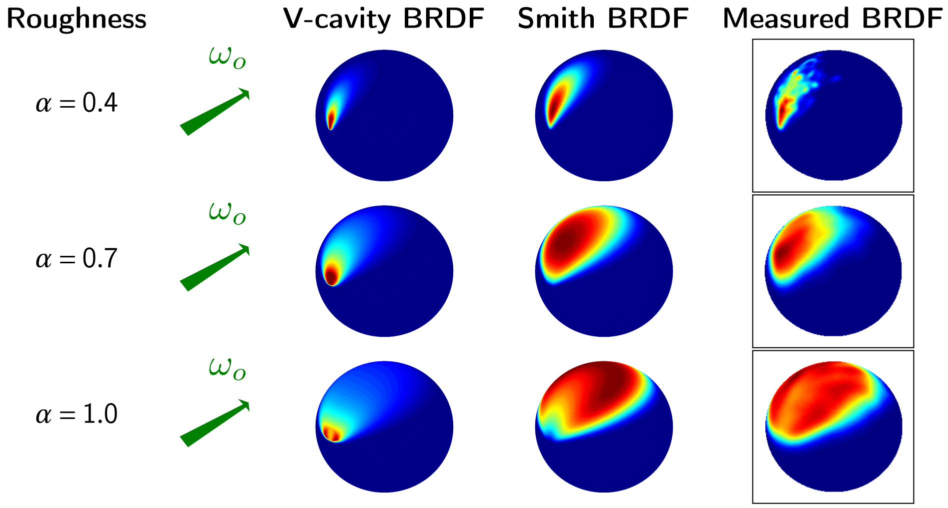
analytic (Beckmann)	microfacet BRDF	BRDF
	$\rho(\omega_o, \omega_i) = \frac{G_2(\omega_o, \omega_i, \omega_h)\, D(\omega_h)}{4\, \cos\theta_o\, \cos\theta_i}$	

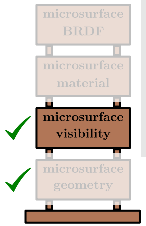


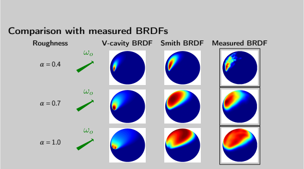


Since the Smith and V-cavity masking functions are both mathematically correct but make different assumptions about the microsurface, we may wonder which one is the most accurate compared to measured data on a continuous, random Gaussian surface. To find out, we can generate such a surface using a noise primitive with Gaussian statistics (Gabor Noise is a good choice). We can then subject this to a raytracing simulation and record the outgoing directions. This gives us a plot of the measured BRDF. Finally, we can compare this against an analytical BRDF model with a compatible Beckmann distribution (parameterized by the Gaussian statistics of the microsurface) and a given masking function.

Comparison with measured BRDFs







We can see that the BRDF predicted by the model with the Smith masking function is much closer to the measured BRDF than the BRDF predicted by the V-cavity masking function.

State of the Art Microfacet Models

State of the Art Microfacet Models

Widely used in production and academia nowadays

- **Beckmann distribution & Smith masking function**
- **GGX distribution & Smith masking function**

Widely used in production and academia nowadays
- Beckmann distribution & Smith masking function
- GGX distribution & Smith masking function

The Beckmann and GGX distributions, with their associated masking functions, are considered state of the art in academia today. They are also the most widely used in the video game and visual effects industries.

State of the Art Microfacet Models

Widely used in production and academia nowadays

- *Isotropic* Beckmann distribution & Smith masking function
- *Isotropic* GGX distribution & Smith masking function

→ Physically based anisotropy?

Widely used in production and academia nowadays

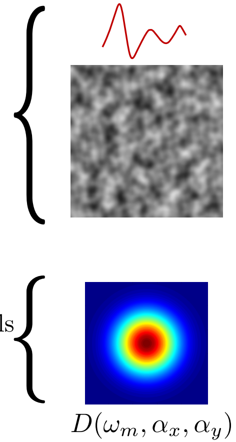
- *Isotropic* Beckmann distribution & Smith masking function
- *Isotropic* GGX distribution & Smith masking function

→ Physically based anisotropy?

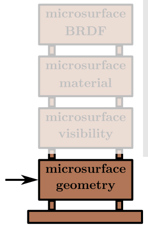
Despite existing for a long time, they were introduced and popularized in CG by Walter et al. in their famous EGSR'07 paper. However, those distributions only model isotropic microsurfaces. It is therefore natural to ask: "can they be extended to anisotropic microsurfaces, whilst retaining their physical properties?"

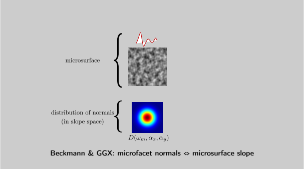
Anisotropy and Stretch Invariance

Anisotropy and Stretch Invariance



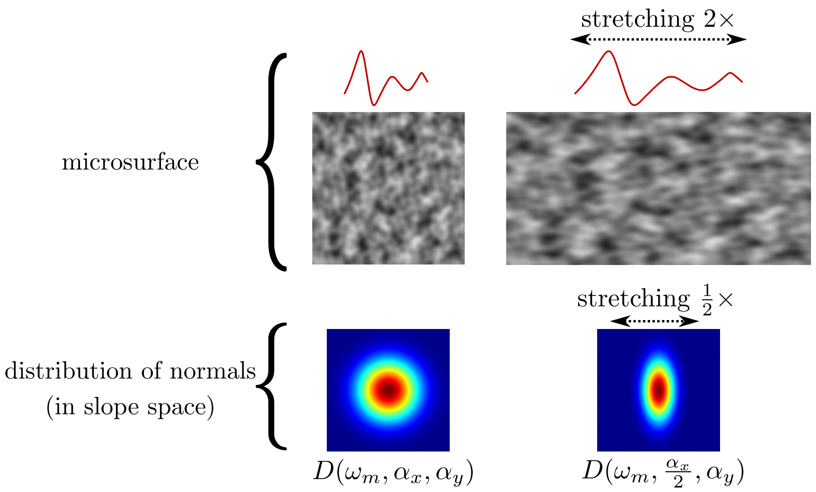
Beckmann & GGX: microfacet normals ⇔ microsurface slope

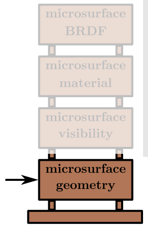




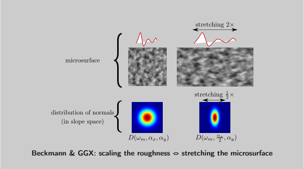
The first step towards anisotropy is to understand the meaning of the roughness parameters α_x and α_y. The Beckmann and GGX distributions have an associated microsurface heightfield, where the normals are given by the slopes of the heightfield.

Anisotropy and Stretch Invariance





Beckmann & GGX: scaling the roughness ⇔ stretching the microsurface



The roughness parameters model how much the heightfield is stretched. For instance, dividing the roughness α_x by a factor of 2 is equivalent to stretching the microsurface by a factor of 2 in the x direction.

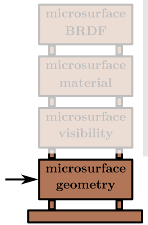
Anisotropy and Stretch Invariance

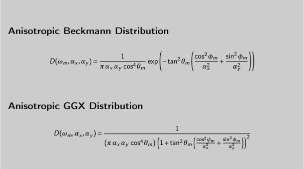
Anisotropic Beckmann Distribution

$$D(\omega_m, \alpha_x, \alpha_y) = \frac{1}{\pi\, \alpha_x\, \alpha_y \cos^4\theta_m} \exp\left(-\tan^2\theta_m \left(\frac{\cos^2\phi_m}{\alpha_x^2} + \frac{\sin^2\phi_m}{\alpha_y^2}\right)\right)$$

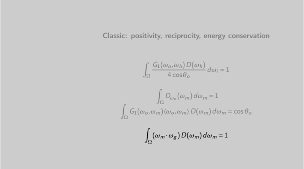
Anisotropic GGX Distribution

$$D(\omega_m, \alpha_x, \alpha_y) = \frac{1}{\left(\pi\, \alpha_x\, \alpha_y \cos^4\theta_m\right)\left(1 + \tan^2\theta_m \left(\frac{\cos^2\phi_m}{\alpha_x^2} + \frac{\sin^2\phi_m}{\alpha_y^2}\right)\right)^2}$$

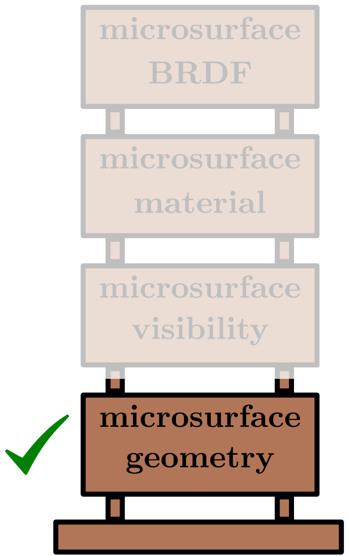




By using this intuition, we can derive anisotropic forms of the Beckmann and GGX distributions.



Validation equations



Classic: positivity, reciprocity, energy conservation

$$\int_\Omega \frac{G_1(\omega_o,\omega_h)\,D(\omega_h)}{4\cos\theta_o}\,d\omega_i = 1$$

$$\int_\Omega D_{\omega_o}(\omega_m)\,d\omega_m = 1$$

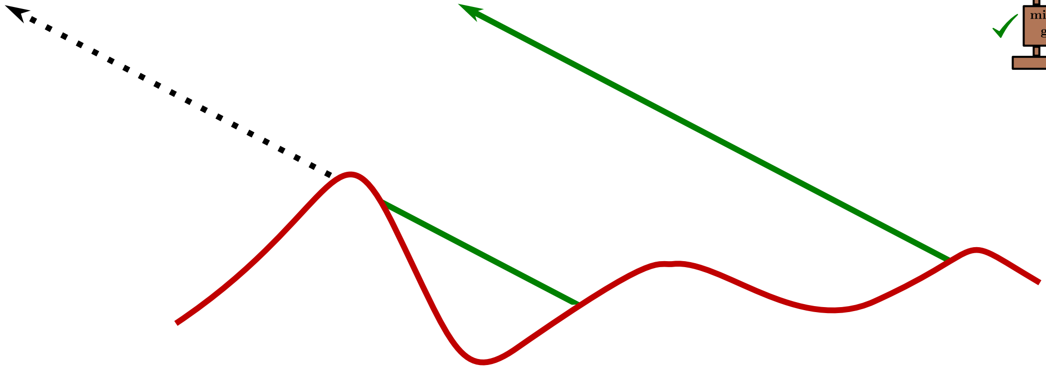
$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\,D(\omega_m)\,d\omega_m = \cos\theta_o$$

$$\int_\Omega (\omega_m\cdot\omega_g)\,D(\omega_m)\,d\omega_m = 1$$

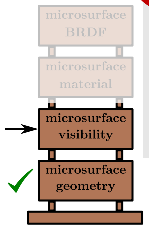
This way of incorporating anisotropy leaves the distributions of normals correctly normalized. What about the next equations related to the masking function?

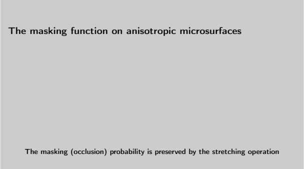
Anisotropy and Stretch Invariance

The masking function on anisotropic microsurfaces



The masking (occlusion) probability is preserved by the stretching operation

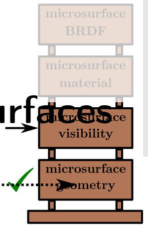




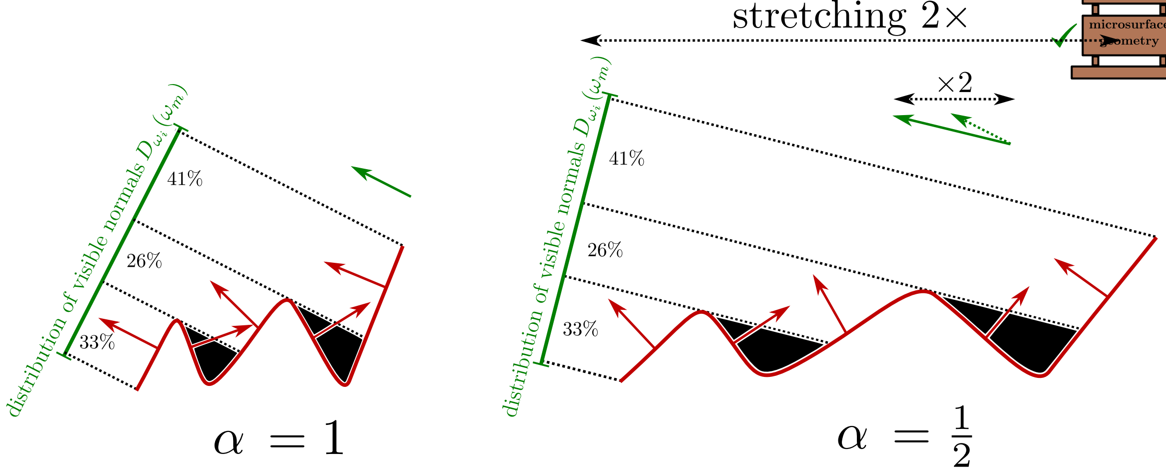
In this animated figure, we can see that after stretching the configuration (the microsurface and the view direction), occluded rays are still occluded and unoccluded ray are still unoccluded. This illustrates an important property: this stretching operation, related to anisotropy, preserves the masking function.

⚠This slide is animated (works with Acrobat Reader).

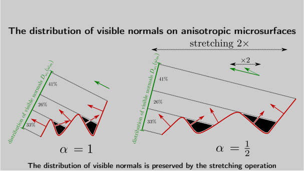
Anisotropy and Stretch Invariance



The distribution of visible normals on anisotropic microsurfaces

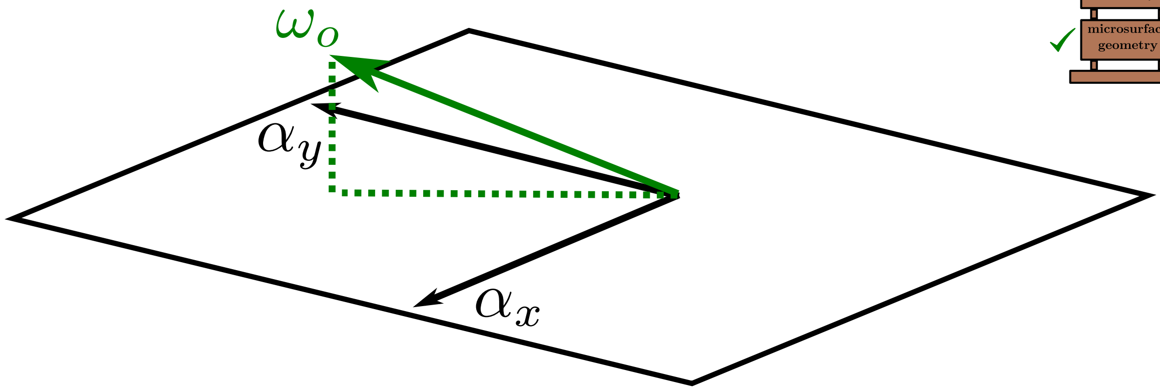


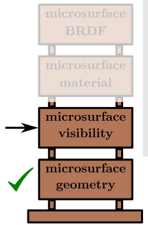
The distribution of visible normals is preserved by the stretching operation



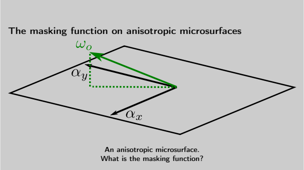
We can see also that the stretching operation preserves the distribution of visible normals.

The masking function on anisotropic microsurfaces



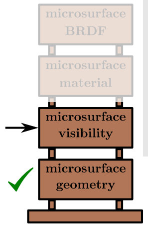


An anisotropic microsurface.
What is the masking function?

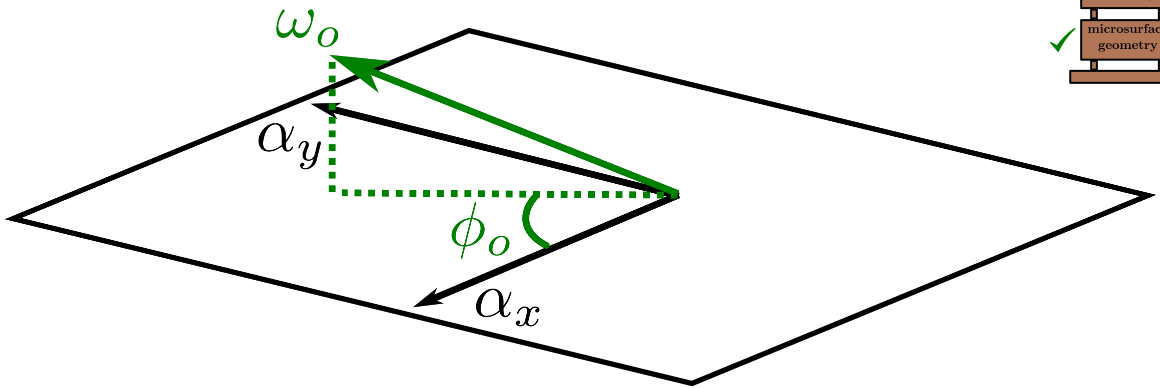


A practical consequence is that if the masking function is known for an isotropic surface, then it is also known for the associated stretched, anisotropic microsurfaces. For instance, this configuration shows a view direction and an anisotropic microsurface...

Anisotropy and Stretch Invariance

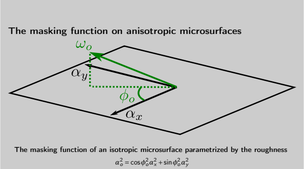


The masking function on anisotropic microsurfaces



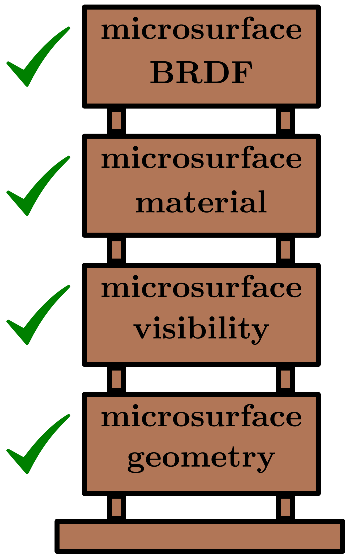
The masking function of an isotropic microsurface parametrized by the roughness

$$\alpha_o^2 = \cos\phi_o^2 \alpha_x^2 + \sin\phi_o^2 \alpha_y^2$$



...and the masking function for this view direction is the masking function of an isotropic microsurface parametrized by the projected roughness α_o.

Validation equations



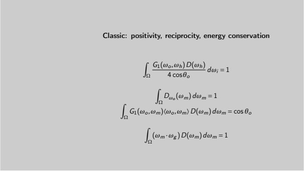
Classic: positivity, reciprocity, energy conservation

$$\int_\Omega \frac{G_1(\omega_o,\omega_h)\, D(\omega_h)}{4\cos\theta_o}\, d\omega_i = 1$$

$$\int_\Omega D_{\omega_o}(\omega_m)\, d\omega_m = 1$$

$$\int_\Omega G_1(\omega_o,\omega_m)\langle\omega_o,\omega_m\rangle\, D(\omega_m)\, d\omega_m = \cos\theta_o$$

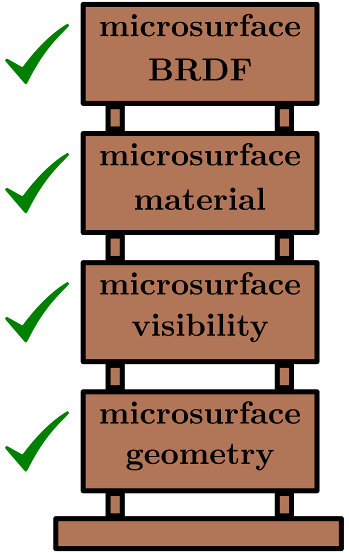
$$\int_\Omega (\omega_m \cdot \omega_g)\, D(\omega_m)\, d\omega_m = 1$$



The masking function derived in this way satisfies the validation equations. It is thus the physically meaningful and correct way to represent anisotropy.

Conclusion

Conclusion

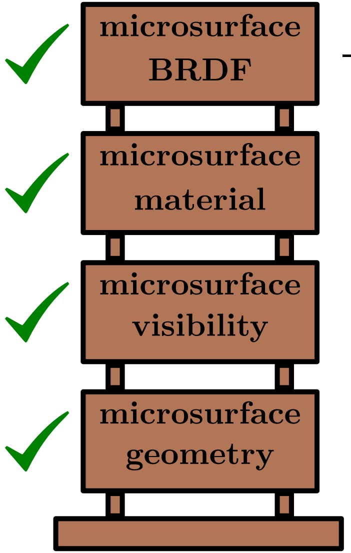


Validation!

Validation!

The purpose of this talk was to provide insights and tools to better understand microfacet models. We have seen that validating microfacet models is important and we have derived strict validation equations to do that. However, it does not mean that the models we call "non-physically based" in this talk shouldn't be used in practice. The goal was simply to emphasize the properties and limitations of the different models on an objective basis, but it is up to you to decide what you want to use in practice.

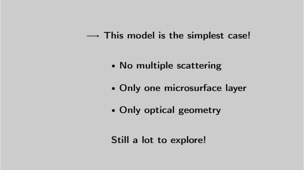
Conclusion



⟶ This model is the simplest case!

- No multiple scattering
- Only one microsurface layer
- Only optical geometry

Still a lot to explore!



As a last remark, I would like to point out that the Cook and Torrance model discussed in this talk, which is also the most widely used one in practice, is actually the simplest case one can think of. It models only single scattering on a one-layer microsurface in the frame of geometric optics. No multiple scattering, no layered materials, no diffraction. Obviously there is still a lot to explore in the realm of shading models, which is also why thorough validation is so important. After all, how can we expect to push the microfacet model further if we can't even get the simplest case right?