Recent Advances in Physically Based Shading

Naty Hoffman
Lucasfilm

Hi. As Steve mentioned, my intro is a bit different this year - covering recent work in the field instead of physics & math fundamentals. And first I want to set expectations regarding...

Scope

Render the Possibilities
SIGGRAPH2016

...scope. This course covers all topics related to physically based shading — from material authoring tools to volume appearance models. Due to time limitations...

Scope

Render the Possibilities
SIGGRAPH2016

...this survey will have a somewhat smaller scope than that.

Shading Models

Render the Possibilities
SIGGRAPH2016

I'll be talking specifically about advances in shading models.

Surface Shading Models

More specifically, surface shading models.

BRDF Surface Shading Models

Even more specifically, BRDF models.

General BRDF Surface Shading Models

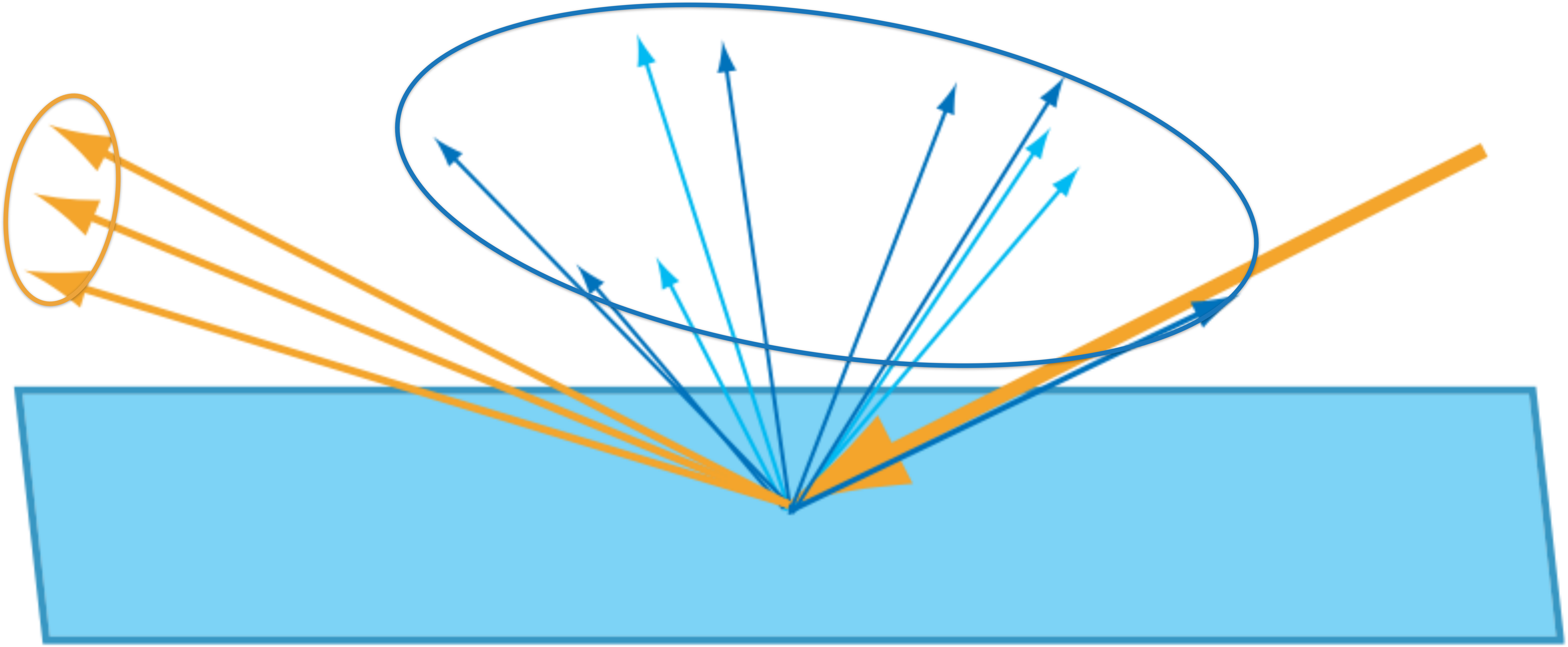
Render the Possibilities
SIGGRAPH2016

And <u>yet</u> more specifically, reasonably general BRDFs, not models that are focused on a specialized class of materials like cloth or car paint or (as we will hear about from Brian Karis later today) hair. Yet even with this modest scope, there will be plenty to talk about.

State of the Art

Before we go into the future with the latest shading model research, let's first take stock of what models are used in the present, in film and game production.

$$f(\mathbf{l},\mathbf{v}) = f_{\text{specular}}(\mathbf{l},\mathbf{v}) + f_{\text{diffuse}}(\mathbf{l},\mathbf{v})$$



Of course, all these models include separate terms for specular and diffuse.

$$f_{\mathrm{specular}}(\mathbf{l}, \mathbf{v}) = \sum_{i=0}^{n} k_i \frac{F_i(\mathbf{l}, \mathbf{h}) G_i(\mathbf{l}, \mathbf{v}, \mathbf{h}) D_i(\mathbf{h})}{4(\mathbf{n} \cdot \mathbf{l})(\mathbf{n} \cdot \mathbf{v})}$$

The specular term is based on microfacet lobes. Games typically only use one lobe, but film shaders often have multiple.

$$\frac{F(\mathbf{l},\mathbf{h})G(\mathbf{l},\mathbf{v},\mathbf{h})D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})}$$

The NDF used is typically...

$$D_{\mathrm{TR}}(\mathbf{h}, \alpha_x, \alpha_y)$$

$$\frac{F(\mathbf{l},\mathbf{h})G(\mathbf{l},\mathbf{v},\mathbf{h})D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})}$$

...Trowbridge-Reitz (also called GGX or GTR2), though some models use Beckmann, typically as an artist-selected option. Here we show an anisotropic NDF with separate x and y roughness parameters. Games tend to rely mainly on isotropic lobes, with anisotropy kept for rare special cases. Film materials tend to use anisotropy more often.

$$\frac{\chi^+(\mathbf{l}\cdot\mathbf{h})\chi^+(\mathbf{v}\cdot\mathbf{h})}{1+\Lambda(\mathbf{l})+\Lambda(\mathbf{v})}$$

$$D_{\mathrm{TR}}(\mathbf{h},\alpha_x,\alpha_y)$$

$$\frac{F(\mathbf{l},\mathbf{h})G(\mathbf{l},\mathbf{v},\mathbf{h})D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})}$$

Current best practice for G is the height-correlated form of Smith, which is more accurate than the uncorrelated form and no more expensive.

$$F_{\text{Schlick}}(F_0, \mathbf{h})$$

$$\frac{\chi^+(\mathbf{l}\cdot\mathbf{h})\chi^+(\mathbf{v}\cdot\mathbf{h})}{1+\Lambda(\mathbf{l})+\Lambda(\mathbf{v})}$$

$$D_{\text{TR}}(\mathbf{h}, \alpha_x, \alpha_y)$$

$$\frac{F(\mathbf{l},\mathbf{h})G(\mathbf{l},\mathbf{v},\mathbf{h})D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})}$$

The Schlick approximation of Fresnel, with its convenient F_0 parameterization, used to be almost universal. And it still is quite common, though there are some other alternatives being used.

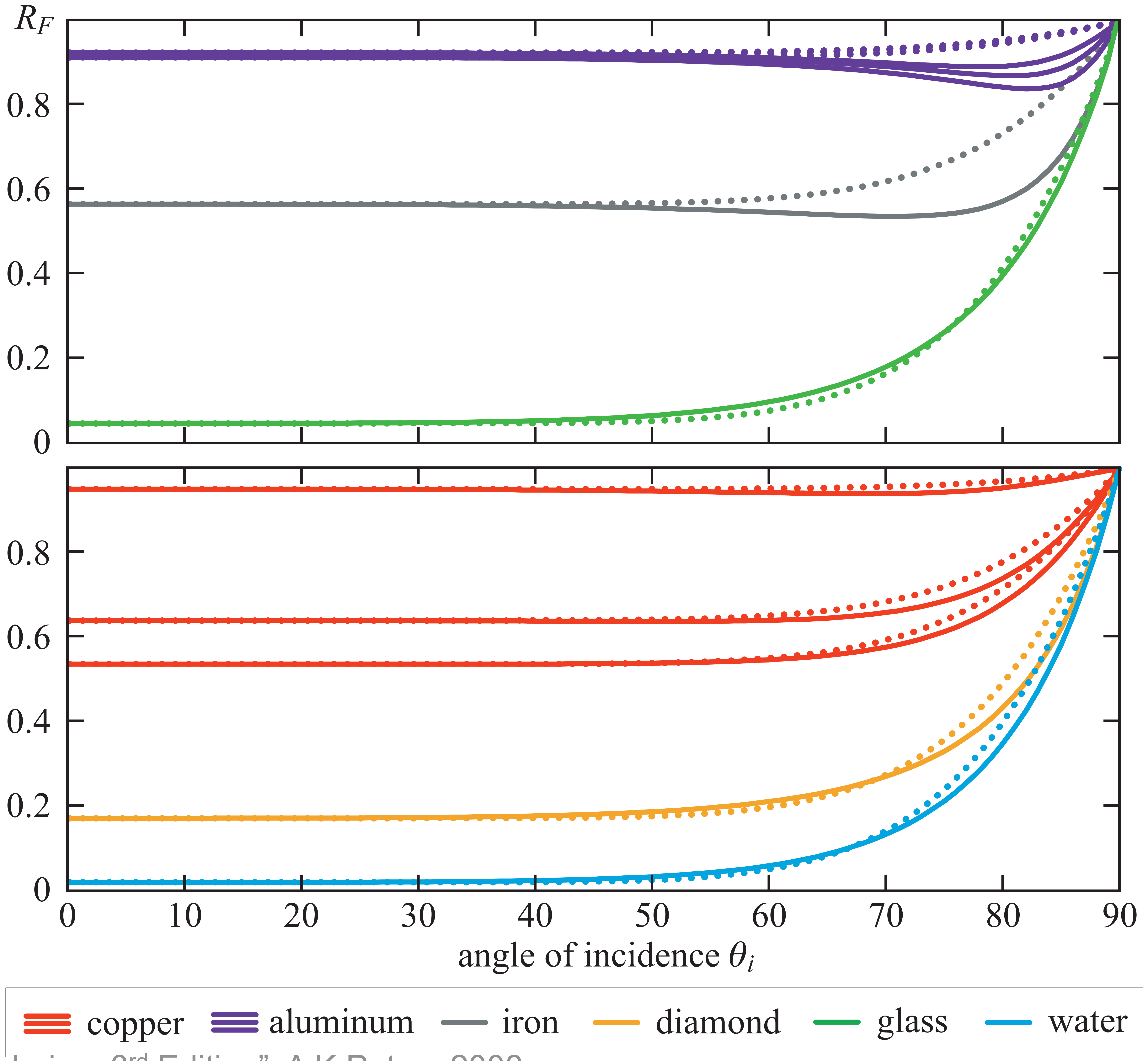


Image from “Real-Time Rendering, 3rd Edition”, A K Peters 2008

While it’s a pretty close match for most materials, for some metals the Schlick approximation (dotted lines in this chart) can diverge a bit at glancing angles from the true Fresnel curves (solid lines). For example we can see that the dielectrics and copper are pretty close, but iron and aluminum are definitely a bit off.

Artist Friendly Metallic Fresnel (JCGT, 2014)

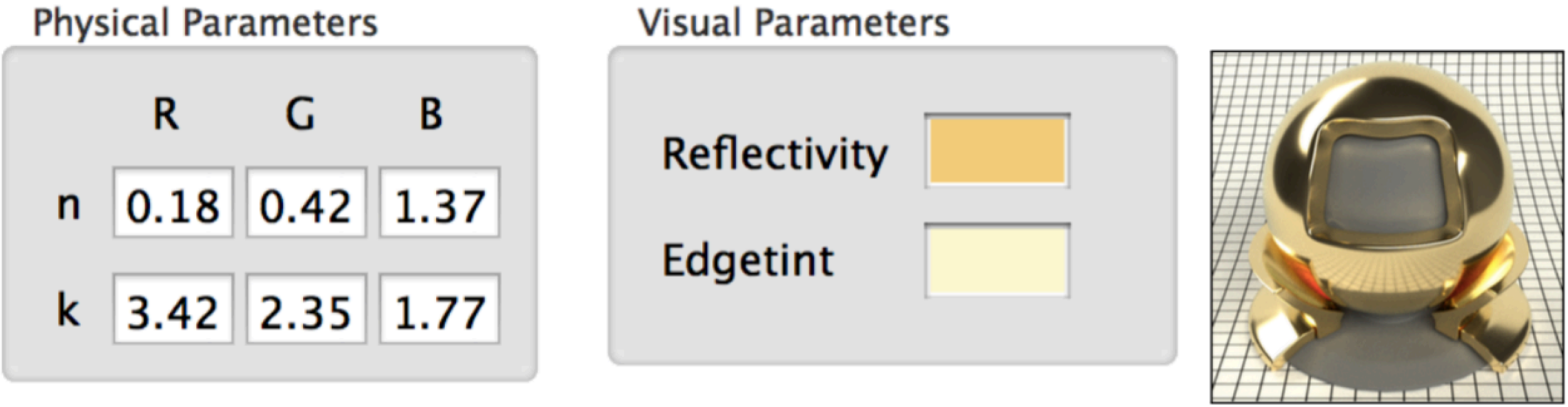


Image from “Artist Friendly Metallic Fresnel”, Gulbrandsen, JCGT 2014

Some film shaders have been using RGB-valued n,k complex index of refraction values to work around that problem. There are a few reasons why RGB values for those quantities is rather nonsensical, but I won’t get into that now. More importantly, they aren’t very intuitive for artists to use, so about two years ago Gulbrandsen from Framestore came up with a more intuitive parameterization, which is now being used in a few different places.

$$f_d = f_{Lambert}\left(1 - 0.5F_L\right)\left(1 - 0.5F_V\right) + f_{retro-reflection}$$

$$f_{Lambert} = \frac{baseColor}{\pi}$$

$$f_{retro-reflection} = \frac{baseColor}{\pi} R_R\left(F_L + F_V + F_L F_V\left(R_R - 1\right)\right)$$

where

$$F_L = \left(1 - \cos\theta_l\right)^5$$

$$F_V = \left(1 - \cos\theta_v\right)^5$$

$$R_R = 2\,roughness\,\cos^2\theta_d$$

Equations from "Extending the Disney BRDF to a BSDF with Integrated Subsurface Scattering", Burley, SIGGRAPH Course Notes 2015

The diffuse BRDF that Brent Burley presented in the course last year is a good example of a state-of-the-art diffuse model; most film shaders use something similar. As do some games, though many still use Lambert.

State of the Art

The current state of the art isn’t too bad overall; people are making some mighty fine-looking movies and games with these.

Issues and Limitations

However, this status quo has some issues and limitations worth discussing.

$$D_{\mathrm{TR}}(\mathbf{h}, \alpha_x, \alpha_y)$$

$$\frac{F(\mathbf{l}, \mathbf{h})G(\mathbf{l}, \mathbf{v}, \mathbf{h})D(\mathbf{h})}{4(\mathbf{n} \cdot \mathbf{l})(\mathbf{n} \cdot \mathbf{v})}$$

Two of these have to do with the available NDFs.

Coarse Microgeometry

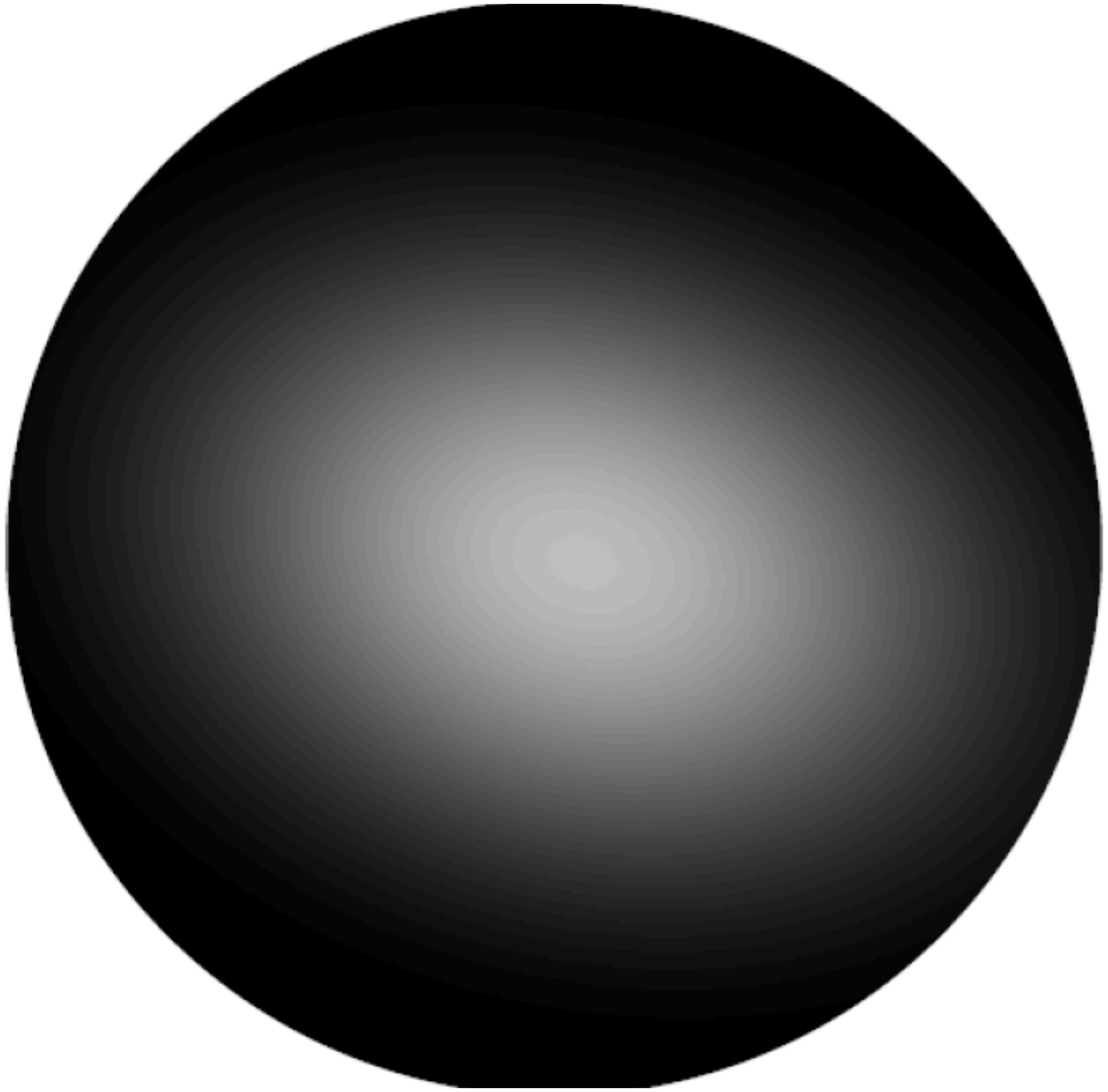


Image from “Rendering Glints on High-Resolution Normal-Mapped Specular Surfaces”, Yan et al., SIGGRAPH 2014

Render the Possibilities
SIGGRAPH2016

The NDFs used today are all smooth, like the one on the left. This is a good representation of very fine-grained microgeometry; with each pixel covering many tens of thousands of surface details. But many surfaces have coarser-grained microstructure, with a pixel only covering perhaps a few dozen surface elements. In that case the NDF looks more like the one on the right, giving the surface a complex “glint” appearance that can’t be reproduced by currently used models.

Shape Control

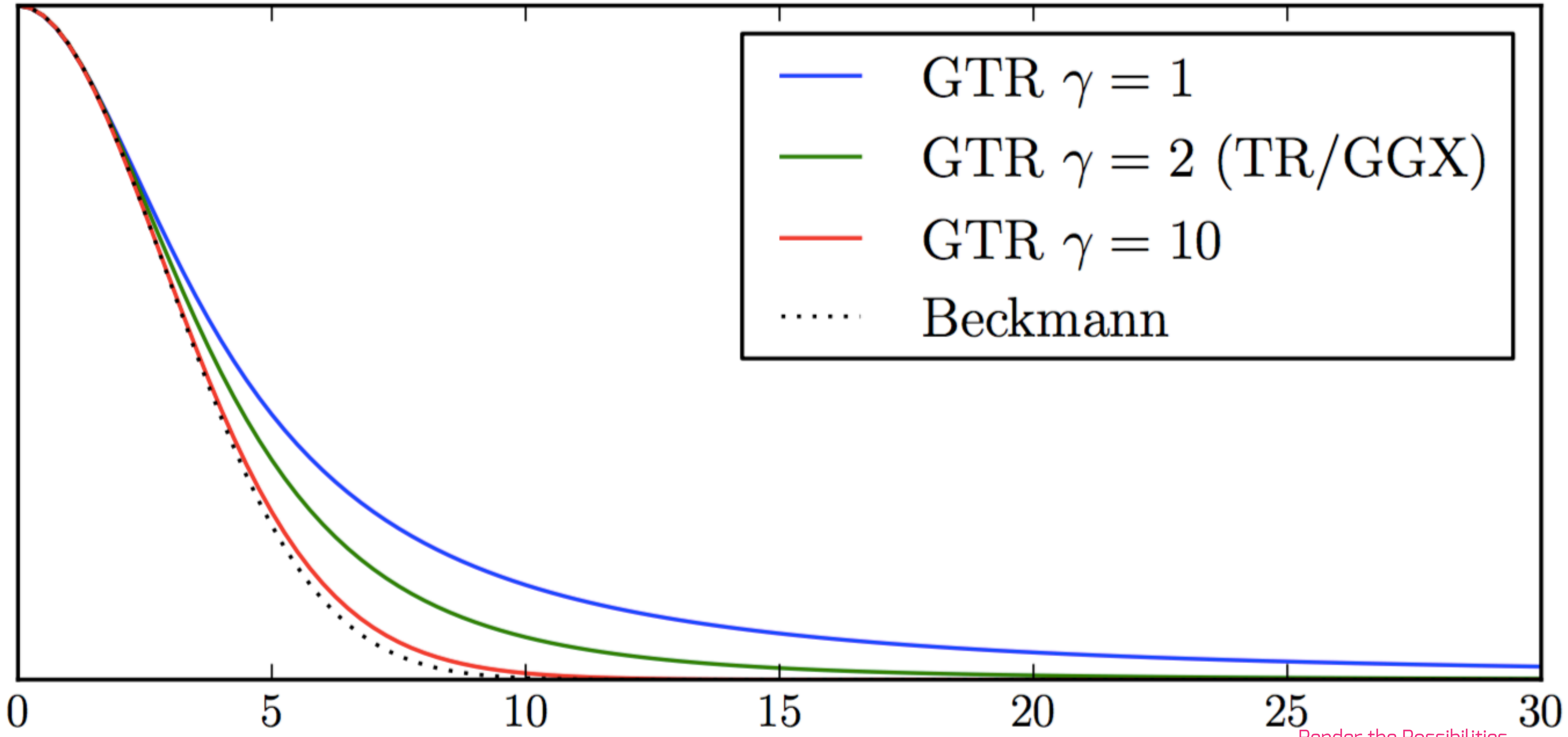


Chart from "Physically Based Shading at Disney", Burley, SIGGRAPH Course Notes 2012

Another issue has to do with shape control. GGX is popular due to its relatively broad "tail", but many materials have broader tails still. It would be great to be able to control that, and in 2012 Brent Burley proposed the Generalized Trowbridge-Reitz (or GTR) NDF to do just that. However, four years have passed and GTR is still not widely used. Why is that?

Shape Invariance

$$D(\theta_m, \alpha) = \frac{f(\frac{\tan\theta_m}{\alpha})}{\alpha^2 \cos^4\theta_m}$$

I believe that the reason has to do with NDF shape invariance. As Eric Heitz showed in his excellent 2014 paper, an isotropic NDF is *shape-invariant* if it has the form shown here, in other words if it can be written as some function of $\tan\theta_m$ over the roughness parameter α, divided by α^2 times $\cos^4\theta_m$. GGX and Beckmann have this form, but GTR does not.

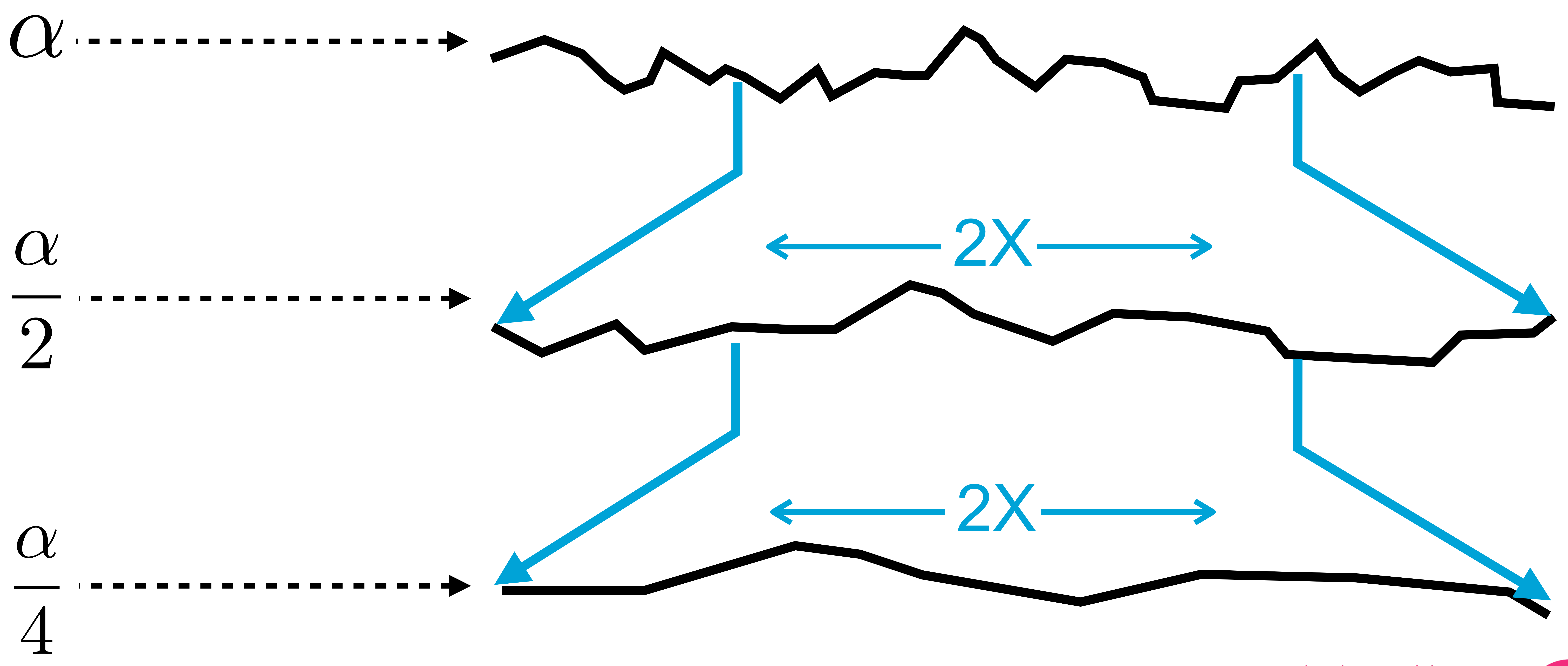
Slope Space

$$P^{22}(x_{\tilde{m}}, y_{\tilde{m}}) = D(\theta_m, \alpha) \cos^4 \theta_m$$

$$P^{22}(x_{\tilde{m}}, y_{\tilde{m}}) = \frac{1}{\alpha^2} f(\frac{\tan \theta_m}{\alpha})$$

This can be more easily understood by looking at the NDF as P^{22}, a distribution of 2D slopes, instead of D, a distribution of 3D vectors. Then we can see that for a distribution with the shape-invariant form, linearly scaling the roughness α causes the distribution in slope space to stretch linearly.

Shape Invariance



And stretching the distribution of slopes is equivalent to stretching the microgeometry, so we see that for shape-invariant NDFs, scaling the roughness parameter is equivalent to stretching the microgeometry by the reciprocal amount.

Shape Invariance - Benefits

- Derivation of normalized anisotropic version
- Derivation of Smith G()
  - One curve (analytic or tabulated) for all roughnesses and anisotropies
- Derivation of importance sampling based on NDF or distribution of visible normals
  - As for Smith G(), a low-dimension function or table handles all roughnesses and anisotropies

Render the Possibilities
SIGGRAPH2016

This property brings many benefits. It makes it much easier to extend the NDF to an anisotropic form, to derive the Smith shadowing-masking function, and to perform importance sampling based on either the NDF or the distribution of visible normals (an effective variance-reducing technique recently introduced by Heitz and D’Eon).

Shape Invariance + Shape Control?

$$D(\theta_m, \alpha, \gamma) = \frac{f(\frac{\tan\theta_m}{\alpha}, \gamma)}{\alpha^2 \cos^4\theta_m}$$

So we want an NDF that (like GTR) has a parameter γ (gamma) to change its shape, but that (unlike GTR) is shape-invariant with respect to the roughness parameter α. But there is no NDF like that in current production use.

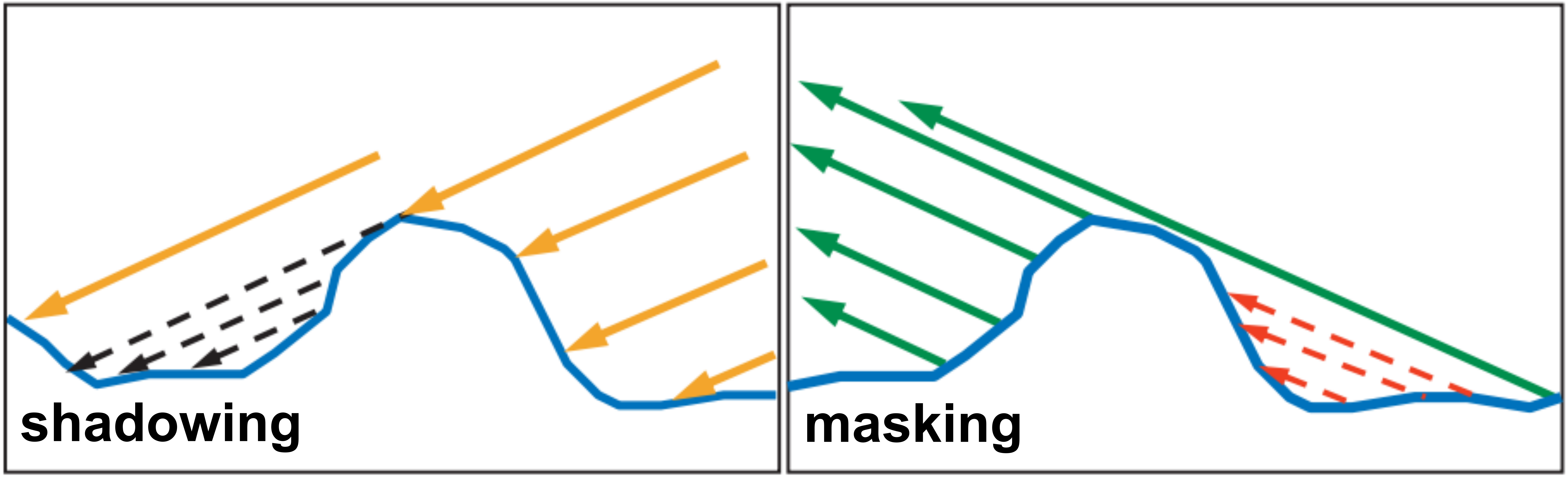
$$\frac{F(\mathbf{l},\mathbf{h})G(\mathbf{l},\mathbf{v},\mathbf{h})D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})}$$

?

Render the Possibilities
SIGGRAPH2016

The other two issues I'll discuss aren't problems with the NDF or some other part of the microfacet BRDF, but basic limitations of microfacet theory itself. Microfacet theory is a very nice and elegant tool which derives BRDF math directly from the surface physical properties — but this very elegance comes from some over-simplified assumptions, which can limit the realism and accuracy of our shading models.

Shadowing and Masking



Images from “Real-Time Rendering, 3rd Edition”, A K Peters 2008

One of these is related to the usage of the G() function, which determines what percentage of otherwise contributing microfacets are occluded from the light or view direction. Which is fine...

Multiple Surface Bounces

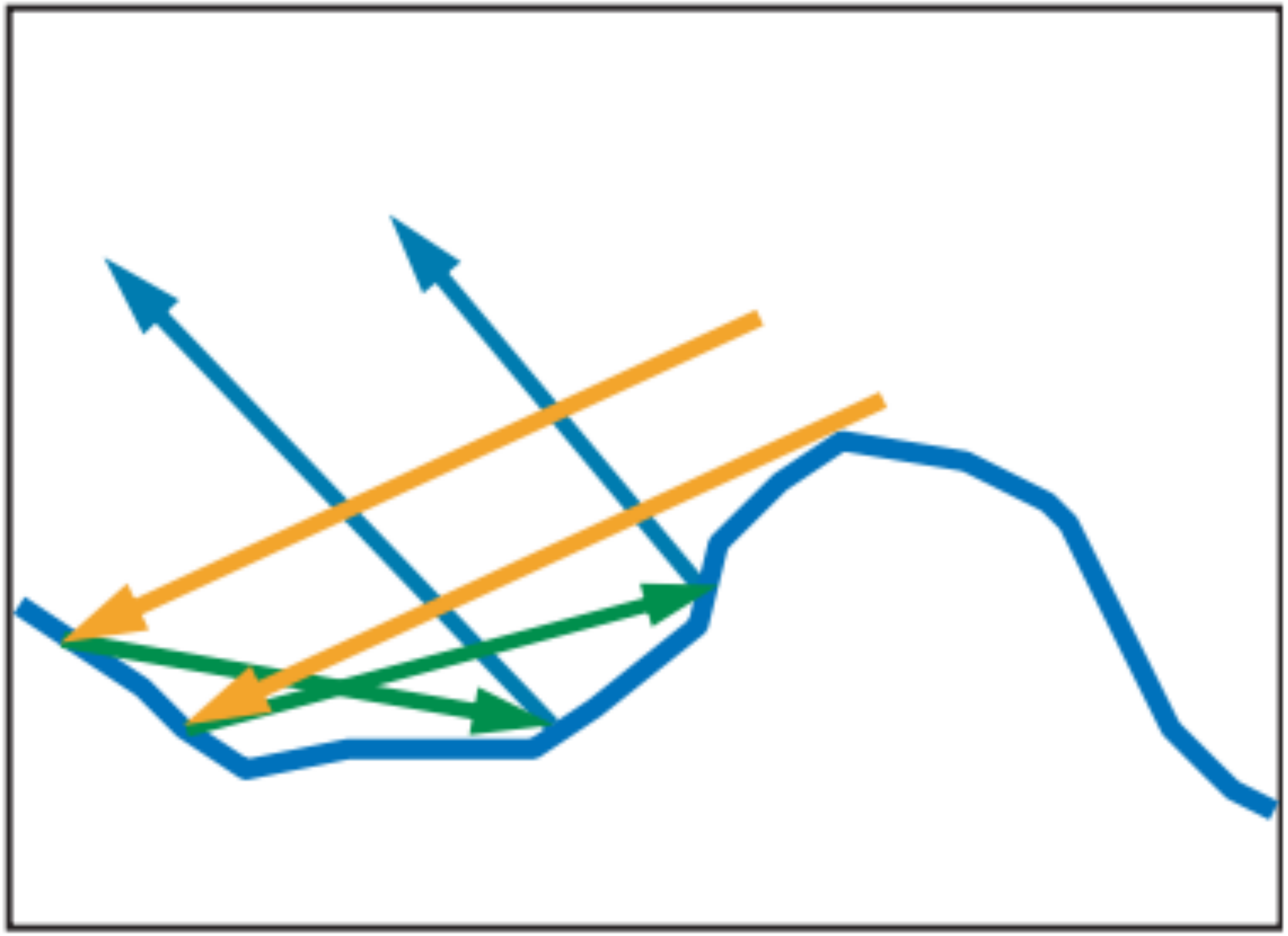


Image from “Real-Time Rendering, 3rd Edition”, A K Peters 2008

...as long as the light that those occluded facets contribute via multiple surface bounce is accounted for.

Multiple Surface Bounces

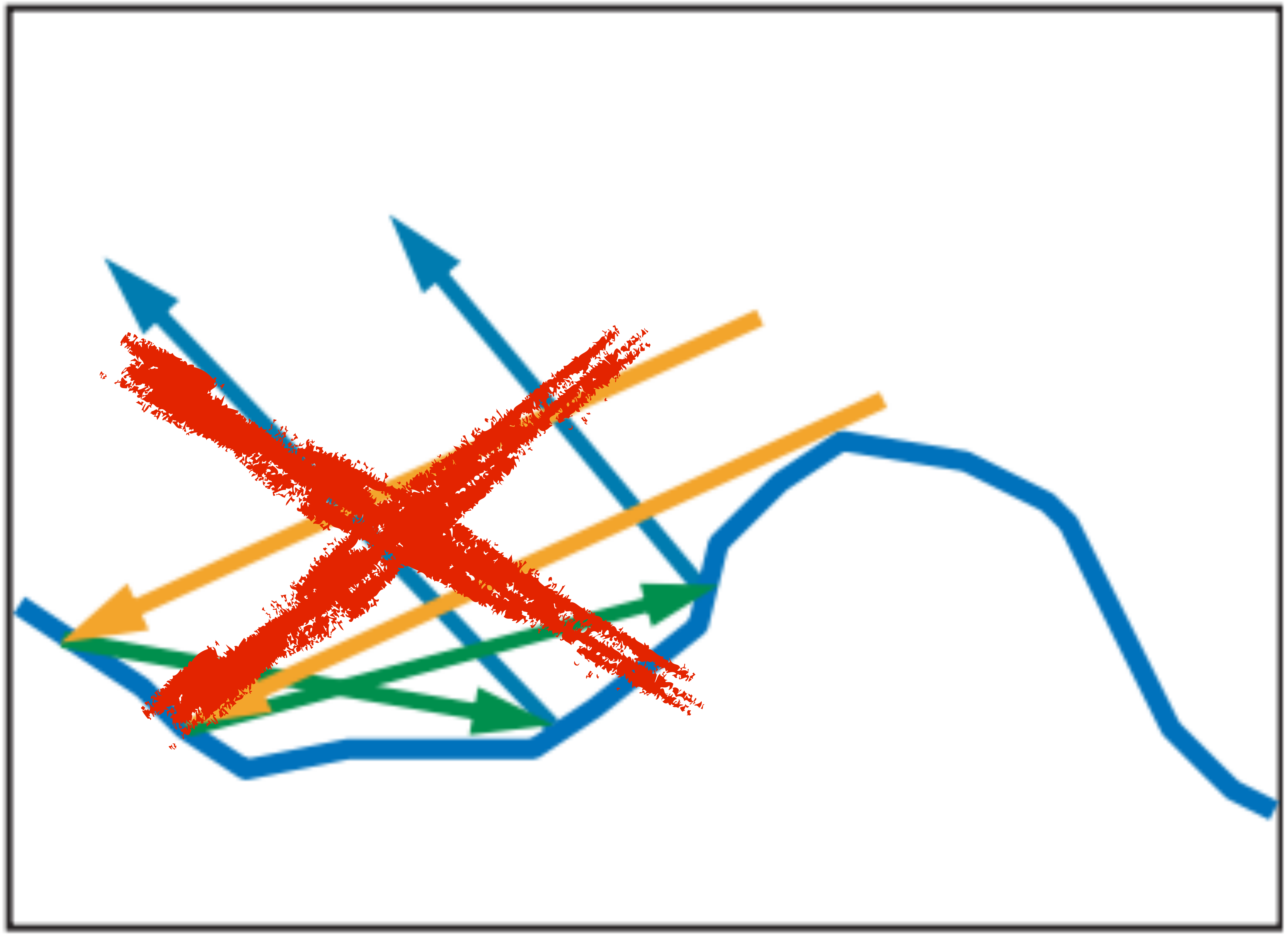
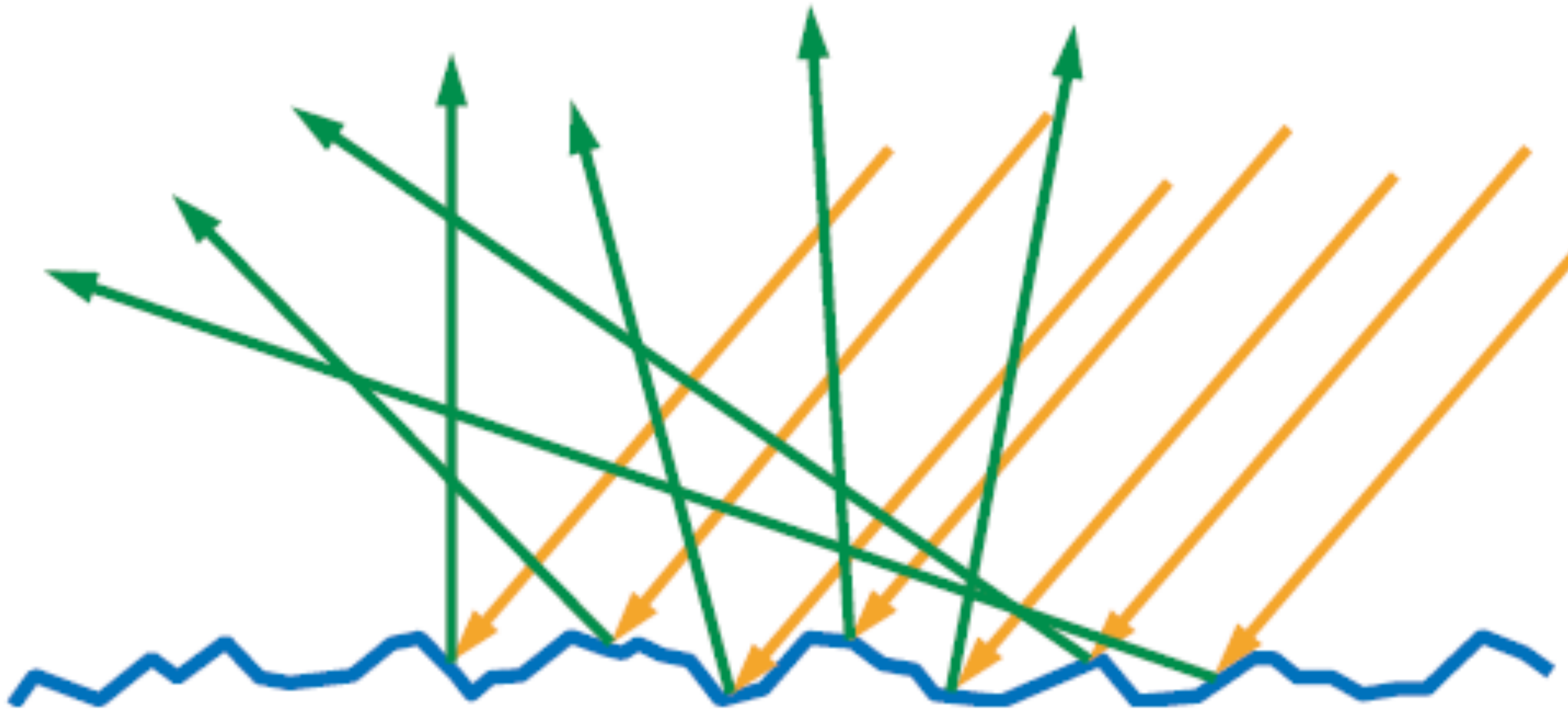


Image from “Real-Time Rendering, 3rd Edition”, A K Peters 2008

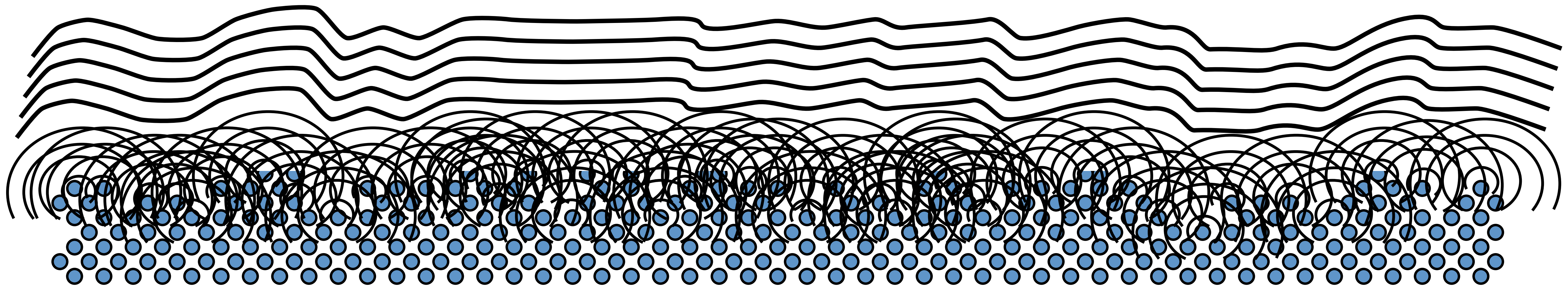
Render the Possibilities
SIGGRAPH2016

But microfacet theory doesn’t account for this, causing an over-darkening due to lost energy. This is sometimes addressed with non-physical correction factors like the Disney model’s “sheen” term.

Geometric Optics



Physical Optics



Render the Possibilities
SIGGRAPH2016

An even more fundamental limitation of microfacet theory is that it's strictly based on ray, or geometric optics, and not the more accurate wave, or physical optics. For many years the rendering community mostly assumed it doesn't matter, but there is recent work indicating that it does. More on this topic in a little bit.

Data-Driven Microfacet Models

Render the Possibilities
SIGGRAPH2016

And now we are finally starting to look at the new stuff. First up, some work on data-driven microfacet models.

Extracting Microfacet-based BRDF Parameters from Arbitrary Materials with Power Iterations (EGSR 2015)

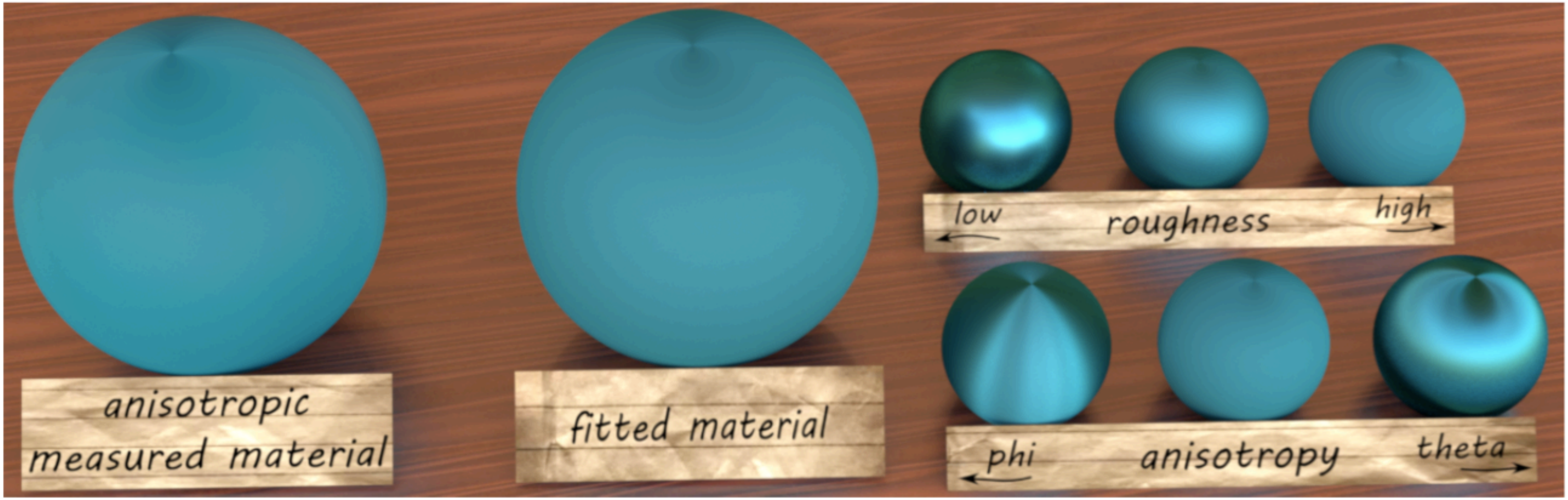
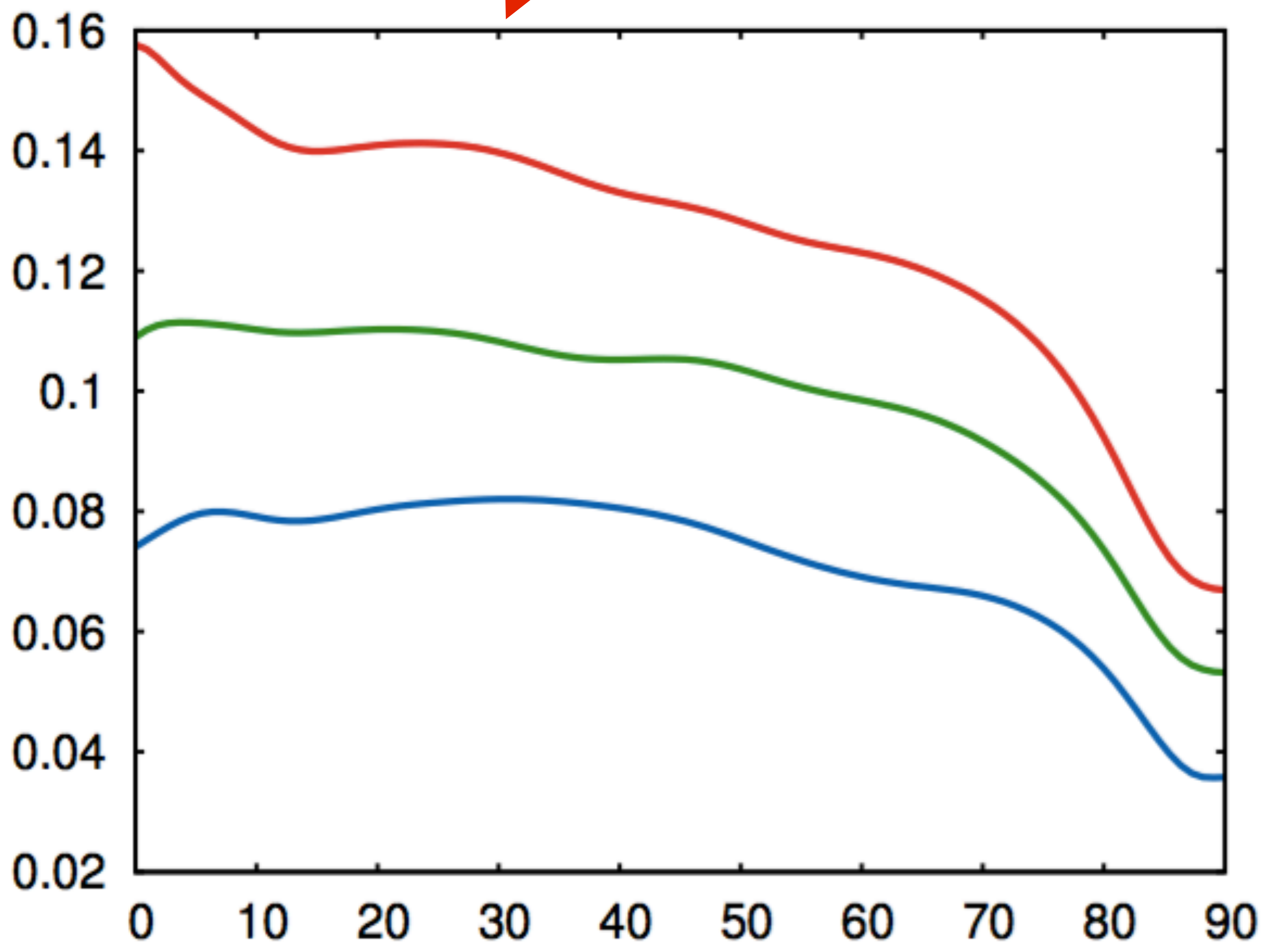
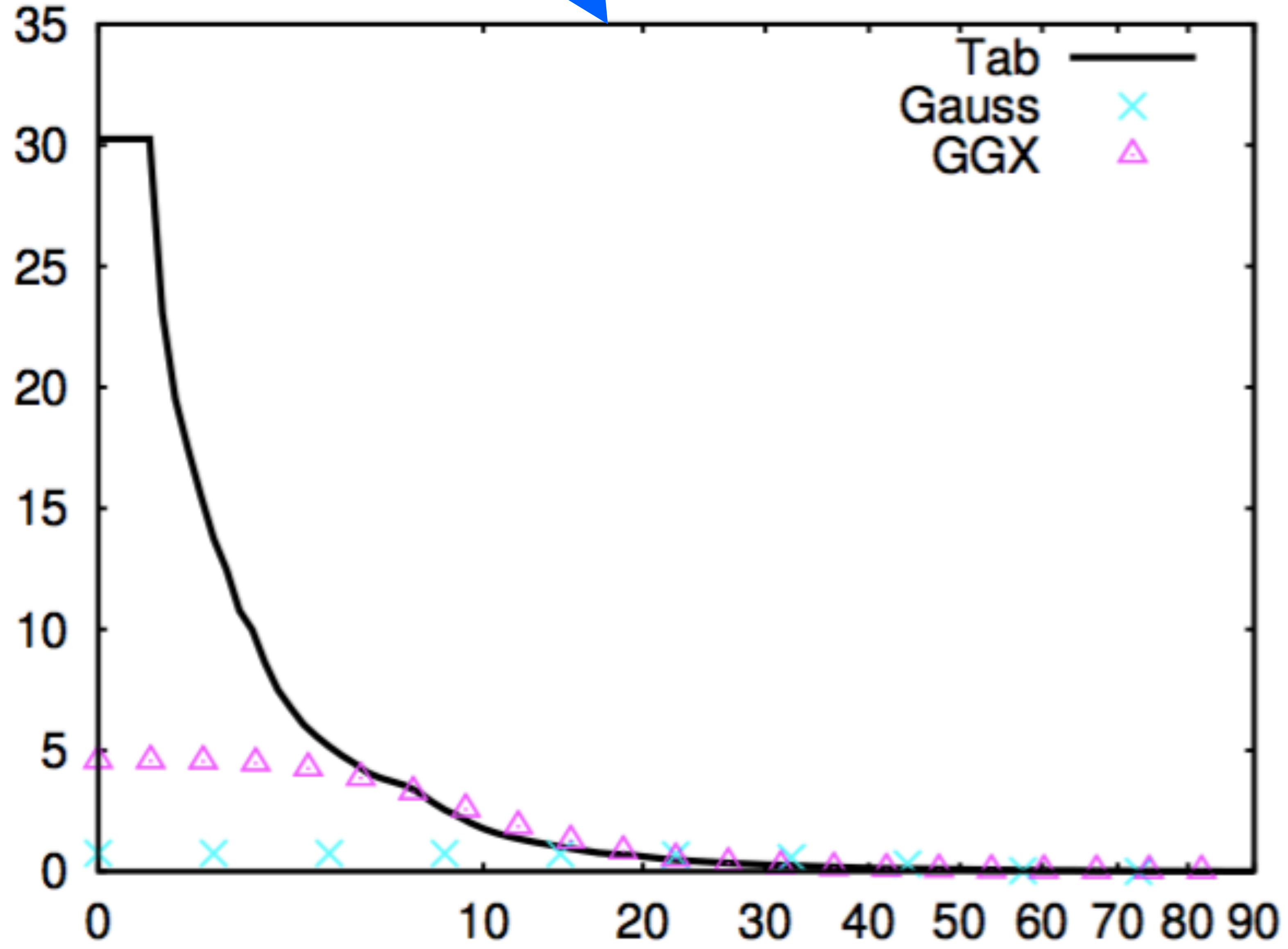


Image from “Extracting Microfacet-based BRDF Parameters from Arbitrary Materials with Power Iterations”, Dupuy et al., EGSR 2015

This paper uses an interesting model for fitting purposes.

$$\frac{F(\mathbf{l},\mathbf{h})G(\mathbf{l},\mathbf{v},\mathbf{h})D(\mathbf{h})}{4(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})}$$





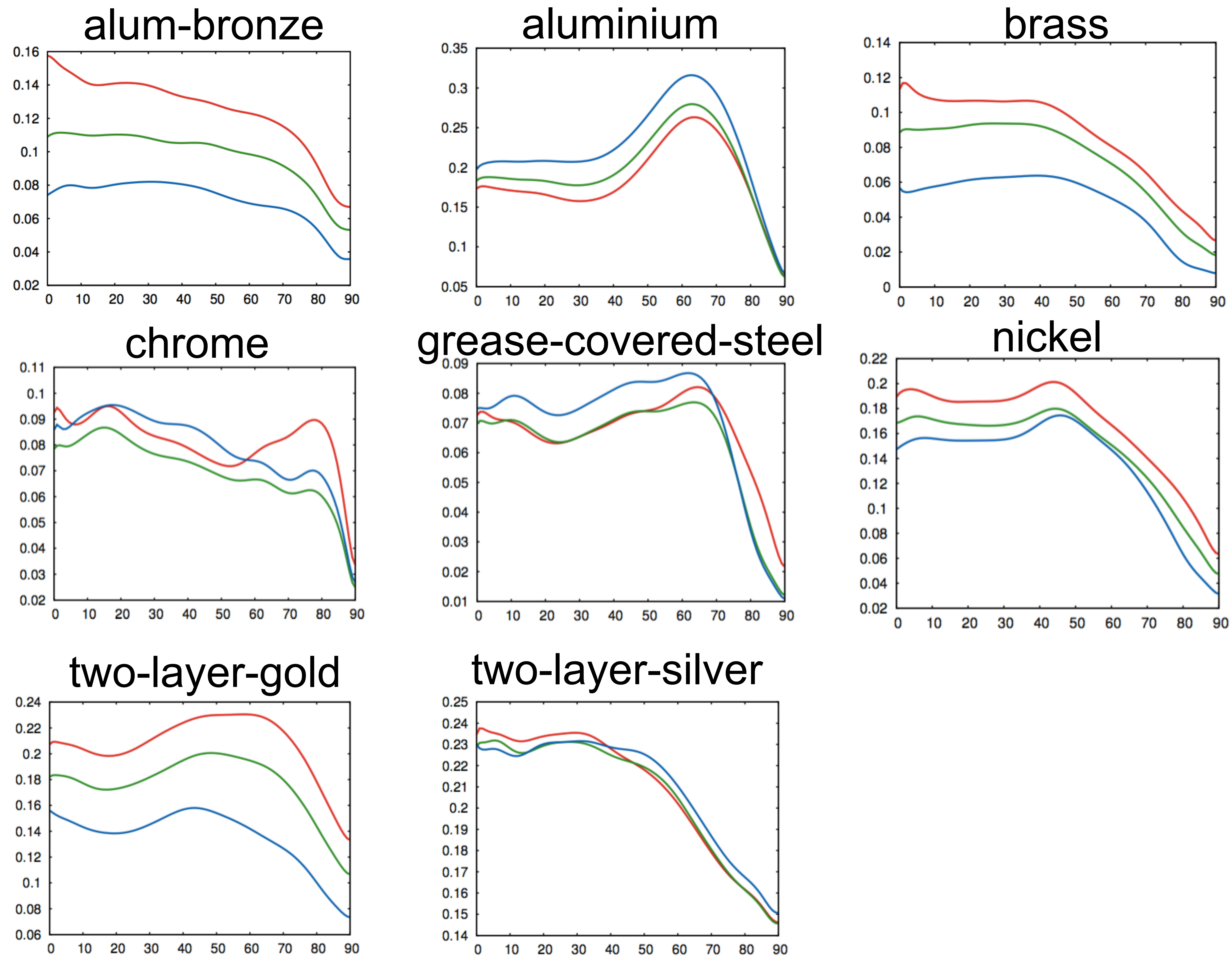
Images from “Extracting Microfacet-based BRDF Parameters from Arbitrary Materials with Power Iterations”, Dupuy et al., EGSR 2015

It can be seen as a middle ground between fully tabulated and fully analytic models. It has the familiar microfacet structure, but D and F are tables extracted from measured BRDF data instead of analytic functions. G is computed as the Smith function based on the extracted D. The NDF is extracted in slope space, which makes it shape-invariant by construction, allowing for varying the roughness and anisotropy. The authors also tried fitting analytic NDFs to the measured ones, and found that GGX produced a reasonable fit in many cases.

Match the theory?

Render the Possibilities
SIGGRAPH2016

So to what extent do the measurements match the predictions of microfacet theory? We know dielectrics won't match well since no diffuse lobe was extracted (an odd choice considering that most of the materials in the MERL database are dielectrics), but the metals are worth looking at. We won't find any contradictions in the extracted NDFs since microfacet theory allows for arbitrary NDFs - but the extracted Fresnel factor can be compared against the expected Fresnel curve for that material.



Images from “Extracting Microfacet-based BRDF Parameters from Arbitrary Materials with Power Iterations”, Dupuy et al., EGSR 2015

Render the Possibilities
SIGGRAPH2016

Here we show the extracted Fresnel curves for all the pure metals (as opposed to oxidized metals or metallic paints) in the MERL database. They look quite different than theory predicts, decreasing towards glancing angles instead of increasing.

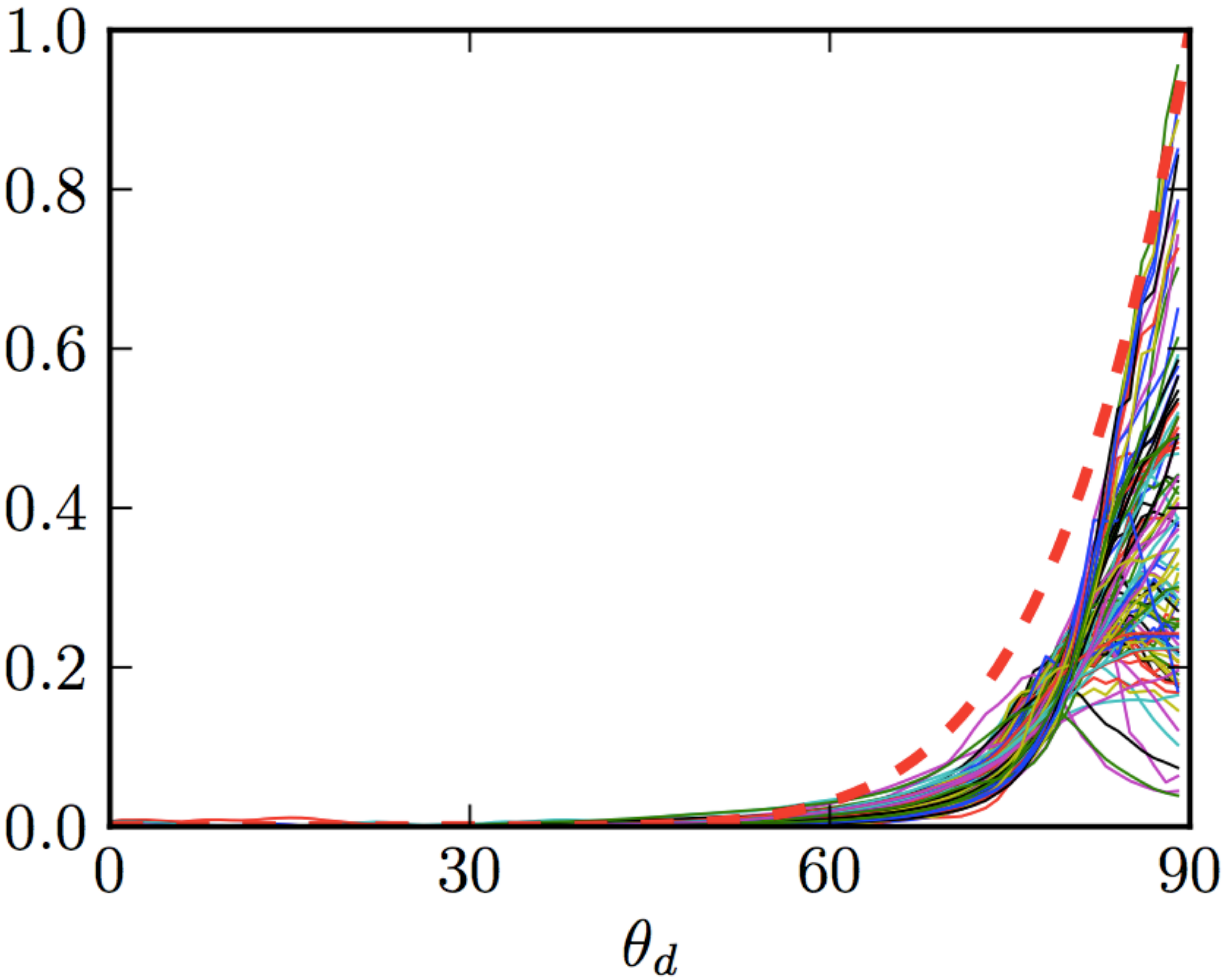
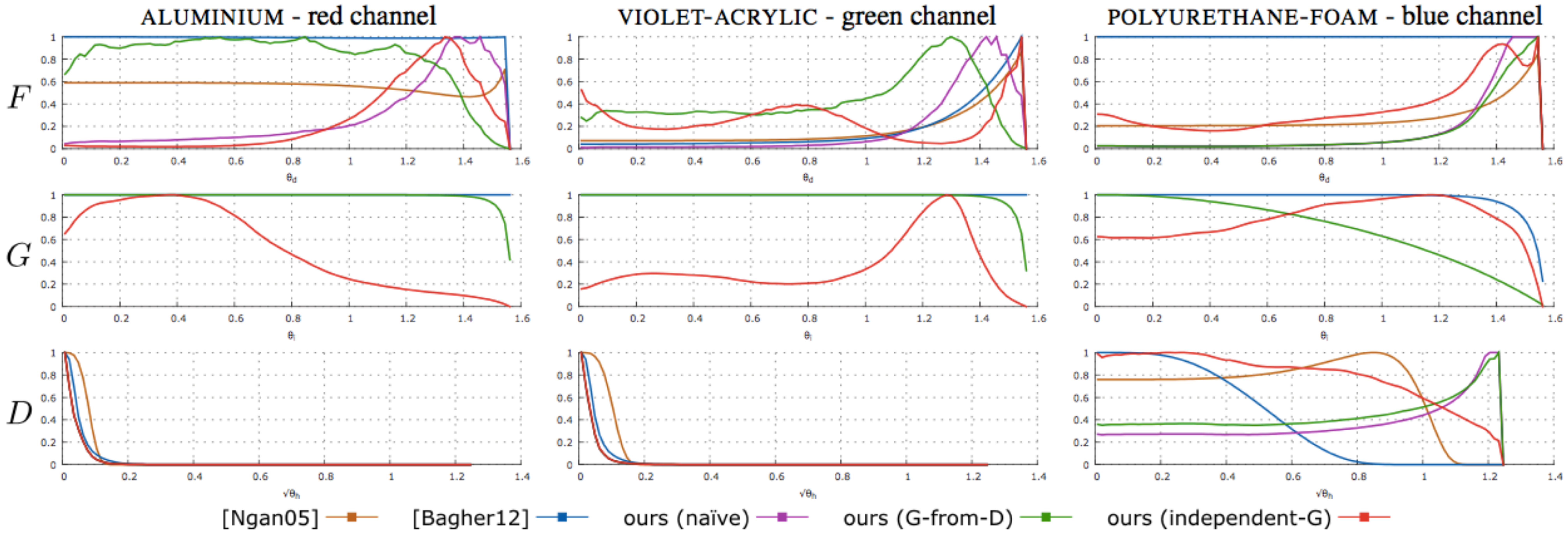
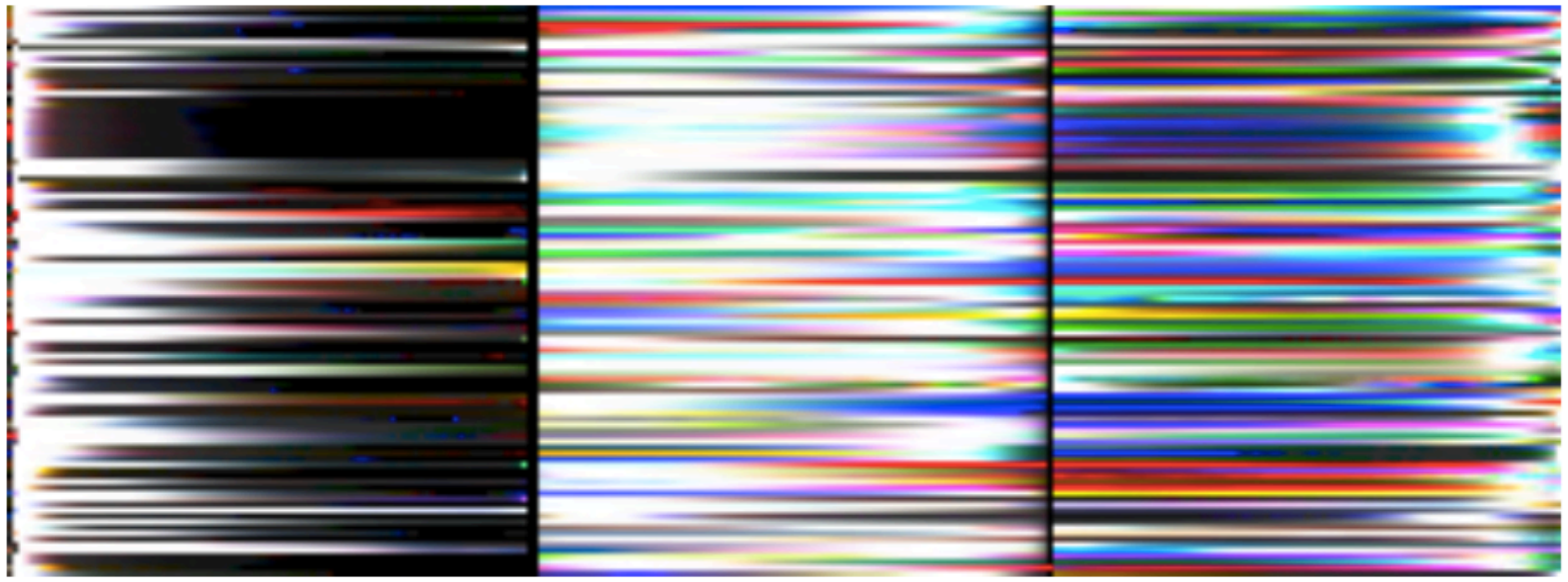


Chart from "Physically Based Shading at Disney", Burley, SIGGRAPH Course Notes 2012

It's interesting to compare to a similar graph that Brent Burley showed in the 2012 course, also taken from the MERL database. This shows an increase for most materials, though some peak around 70 degrees and those could be the same materials as the previous graph. Worth a closer look.

A Non-Parametric Factor Microfacet Model for Isotropic BRDFs (SIGGRAPH 2016)





Images from “A Non-Parametric Factor Microfacet Model for Isotropic BRDFs”, Bagher et al., SIGGRAPH 2016

Render the Possibilities
SIGGRAPH2016

This paper covers some similar ground to the last one, but with some important differences. It uses more of an optimization or fitting approach, is restricted to isotropic BRDFs, and it has a Lambertian term in addition to the microfacet one.

$$\rho_d + \rho_s \frac{F(\mathbf{l}, \mathbf{h}) G(\mathbf{l}, \mathbf{h}) G(\mathbf{v}, \mathbf{h}) D(\mathbf{h})}{(\mathbf{n} \cdot \mathbf{l})(\mathbf{n} \cdot \mathbf{v})}$$

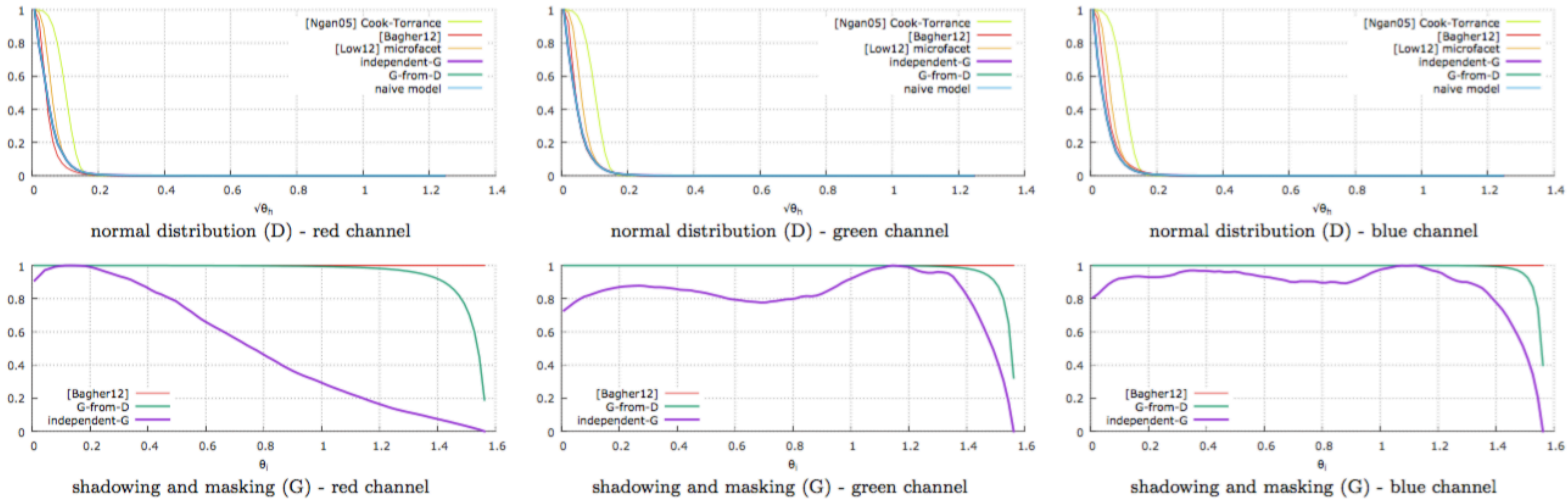
More importantly, their model is much less constrained by microfacet theory than that of the previous paper. They fit separate D and G functions per color channel, and they have a specular coefficient that doesn't match any aspect of the theoretical model. Regarding G, they tried two options: deriving it from D using the generalized Smith approach (like the last paper), and fitting an independent curve.

Match the theory?

It’s worth asking the same question.

Chart from “A Non-Parametric Factor Microfacet Model for Isotropic BRDFs”, Bagher et al., SIGGRAPH 2016

The authors claimed that the independent G (red) did much better than Smith G (light green), especially for more diffuse materials (which are to the right in this graph). If true, that would be a point against the accuracy of Smith G. The difference in accuracy didn’t seem that striking to me when I looked at the data, but I only had time for a relatively quick look.



Charts from “A Non-Parametric Factor Microfacet Model for Isotropic BRDFs”, Bagher et al., SIGGRAPH 2016

Another potential contradiction with microfacet theory would be if the color channels of the G or D function differed significantly. From my quick look, they seemed to be pretty close except for some structured materials where such differences would not be unexpected (one of them - metallic red paint - is shown on the slide). Overall, I’d say the data is worth a closer look.

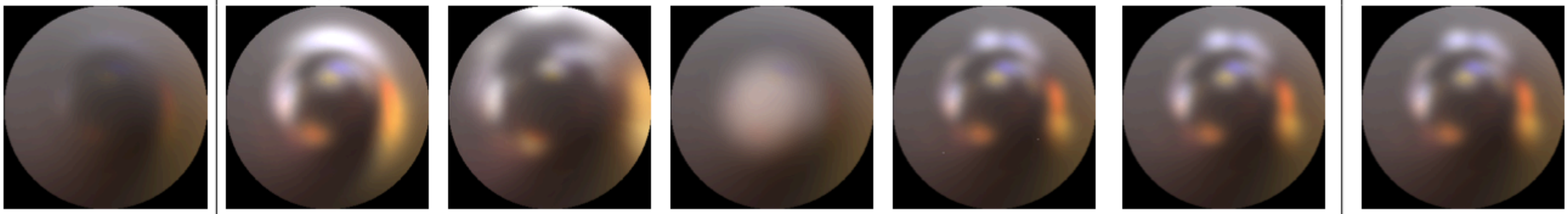
Analytic Microfacet Models

Render the Possibilities
SIGGRAPH2016

We'll now go over some recently published analytic microfacet models.

genBRDF: Discovering New Analytic BRDFs with Genetic Programming (SIGGRAPH 2014)

Initial BRDF	Generation 1	Generation 2	Generation 5	Generation 6	Generation 10	Target BRDF
$n = 4.55$, Err.: 5.12	$n = 15.1$, Err.: 4.83	$n = 0.27$, Err.: 8.10	$n = 1.16$, Err.: 5.72	$n = 80.9$, Err.: 0.077	$n = 82.0$, Err.: 0.0	$n' = 80.0$



$\cos(R \cdot V)^n$	$n\ \cos(R \cdot V)^n$	$n\ \cos(R \cdot H)^n$	$\frac{n}{2\pi} \cos(N \cdot V)^n$	$\frac{n}{2\pi} \cos(N \cdot H)^n$	$\frac{n}{2\pi} \cos(N \cdot H)^{n-2}$	$\frac{n'+2}{2\pi} \cos(N \cdot H)^{n'}$
	Insert(n)	Insert(n) Replace(V,H)	Insert(n) Replace(R, N) Insert($\frac{}{2\pi}$)	Insert(n) Replace(V, H) Replace(R, N) Insert($\frac{}{2\pi}$)	Insert(n) Replace(V, H) Replace(R, N) Insert($\frac{}{2\pi}$) Insert($^{-2}$)	

Image from “genBRDF: Discovering New Analytic BRDFs with Genetic Programming”, Brady et al., SIGGRAPH 2014

Render the Possibilities
SIGGRAPH2016

This paper took an interesting approach to finding analytic models that were a good fit for measured materials — genetic programming. They had a few promising results, at least one of which was further tweaked and turned into a model we’ll discuss in a little bit.

Shape Invariance + Shape Control?

$$D(\theta_m, \alpha, \gamma) = \frac{f(\frac{\tan\theta_m}{\alpha}, \gamma)}{\alpha^2 \cos^4\theta_m}$$

Render the Possibilities
SIGGRAPH2016

I mentioned earlier the need for an NDF that combines shape control and shape invariance. Well, while doing research for this talk I ran into two candidates. I haven't had a chance to try them out much, so I'm putting them out there for other people to experiment with.

NDF: Generalized Beckmann

$$D_{\mathrm{GB}}(\theta) = \frac{\gamma}{\pi\alpha^2\Gamma(1/\gamma)\cos^4\theta} e^{-\left(\frac{\tan^2\theta}{\alpha^2}\right)^\gamma}$$

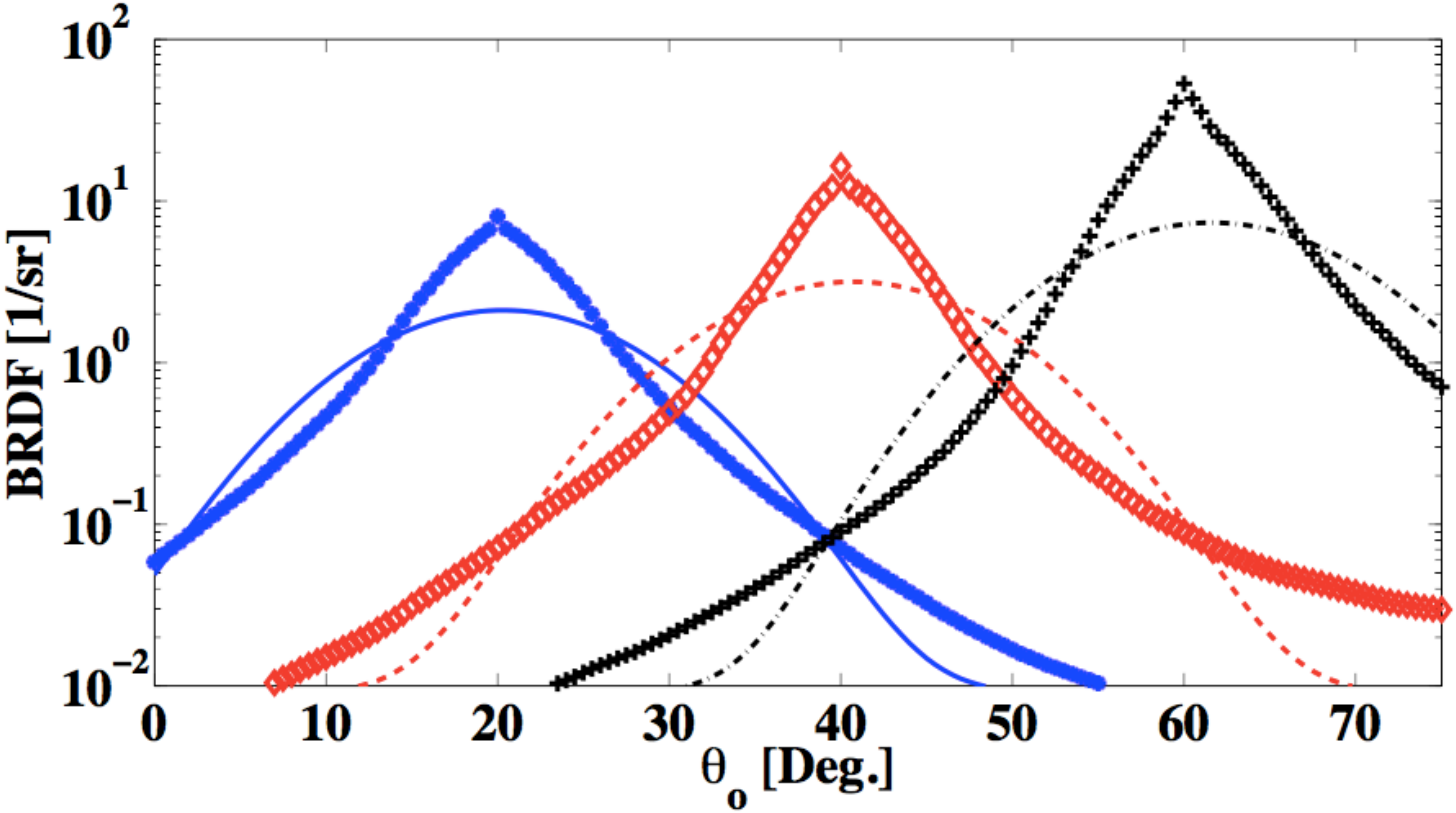
Render the Possibilities
SIGGRAPH2016

Generalized Beckmann is based on one of the results from the genBRDF paper I mentioned earlier, though the shape-invariance was added later. It's a secondary contribution in an upcoming white paper by Holzschuch and Pacanowski — I'll be discussing the paper a bit later, so for now let's focus on the NDF. It has a shape parameter γ, and a roughness parameter α — when γ is equal to 1 then it's identical to Beckmann. Like regular Beckmann, Generalized Beckmann is shape-invariant with respect to α. The γ parameter controls the kurtosis of the distribution - smaller values give it a spikier peak and a long, even lumpy tail.

NDF: Hyper-Cauchy

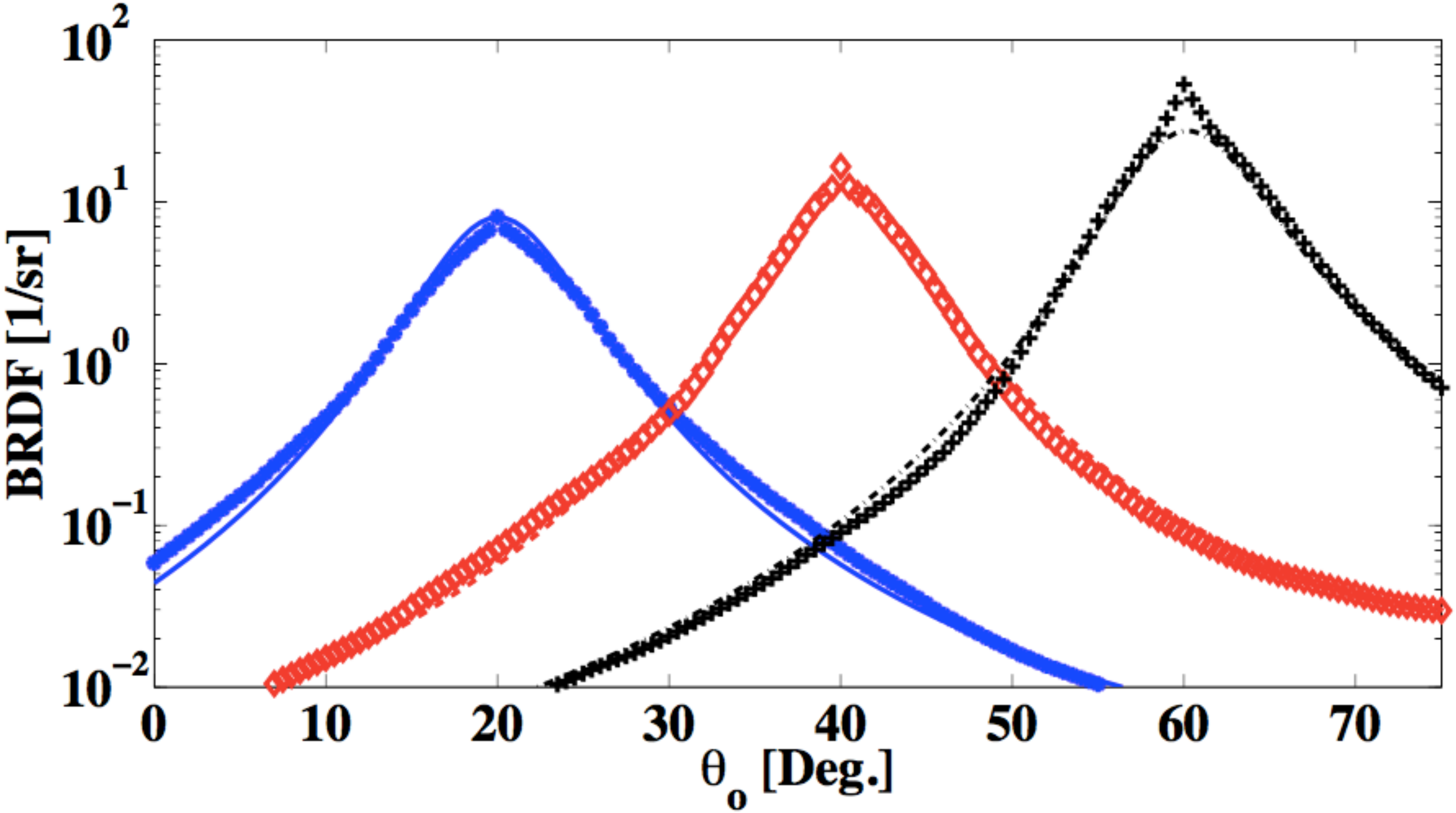
$$D_{\mathrm{HC}}(\theta) = \frac{(\gamma - 1)\left(\sqrt{2}\right)^{2\gamma - 2}}{\pi \alpha^2 \cos^4 \theta \left(2 + \frac{\tan^2 \theta}{\alpha^2}\right)^{\gamma}}$$

The Hyper-Cauchy distribution originated in a 2006 paper by Wellems et al., but it came to my attention in a paper by Butler, who proposed using it to fit measured BRDFs. Like generalized Beckmann, it has a shape control parameter, and a roughness parameter, to which it is shape-invariant.



Charts from “Robust Categorization of Microfacet BRDF Models to Enable Flexible Application-specific BRDF Adaptation”, Butler & Marciniak, Proc. SPIE 9205, Reflection, Scattering, and Diffraction from Surfaces IV, (2014)

And it looks like Butler knew what he was talking about. Here’s an attempt to fit the MERL Nickel BRDF with Beckmann. The three colors are three different incidence angles, the symbols are the measured data and the dashed or solid lines are the fit.



Charts from "Robust Categorization of Microfacet BRDF Models to Enable Flexible Application-specific BRDF Adaptation", Butler & Marciniak, Proc. SPIE 9205, Reflection, Scattering, and Diffraction from Surfaces IV, (2014)

Render the Possibilities
SIGGRAPH2016

And here's a fit with Hyper-Cauchy. A lot better. Another interesting fact about Hyper-Cauchy: when setting γ to a value of 2, it becomes very similar to GGX (at least in terms of the math — I haven't compared the curves yet). So maybe this could be a good replacement for GTR.

Microflakes and Multiple Surface Scattering

The next part of the talk will cover microflake theory and multiple surface scattering.

Microflakes

Image from "A Radiative Transfer Framework for Rendering Materials with Anisotropic Structure", Jakob et al., SIGGRAPH 2010

Microflakes were first introduced in 2010, to model volumes with anisotropic structure, such as cloth and fibrous tissue. They are an extension of the concept of scattering particles, which may have an anisotropic phase function but are themselves isotropic in the sense that the properties of the medium are invariant to rotation.

The SGGX Microflake Distribution (SIGGRAPH 2015)

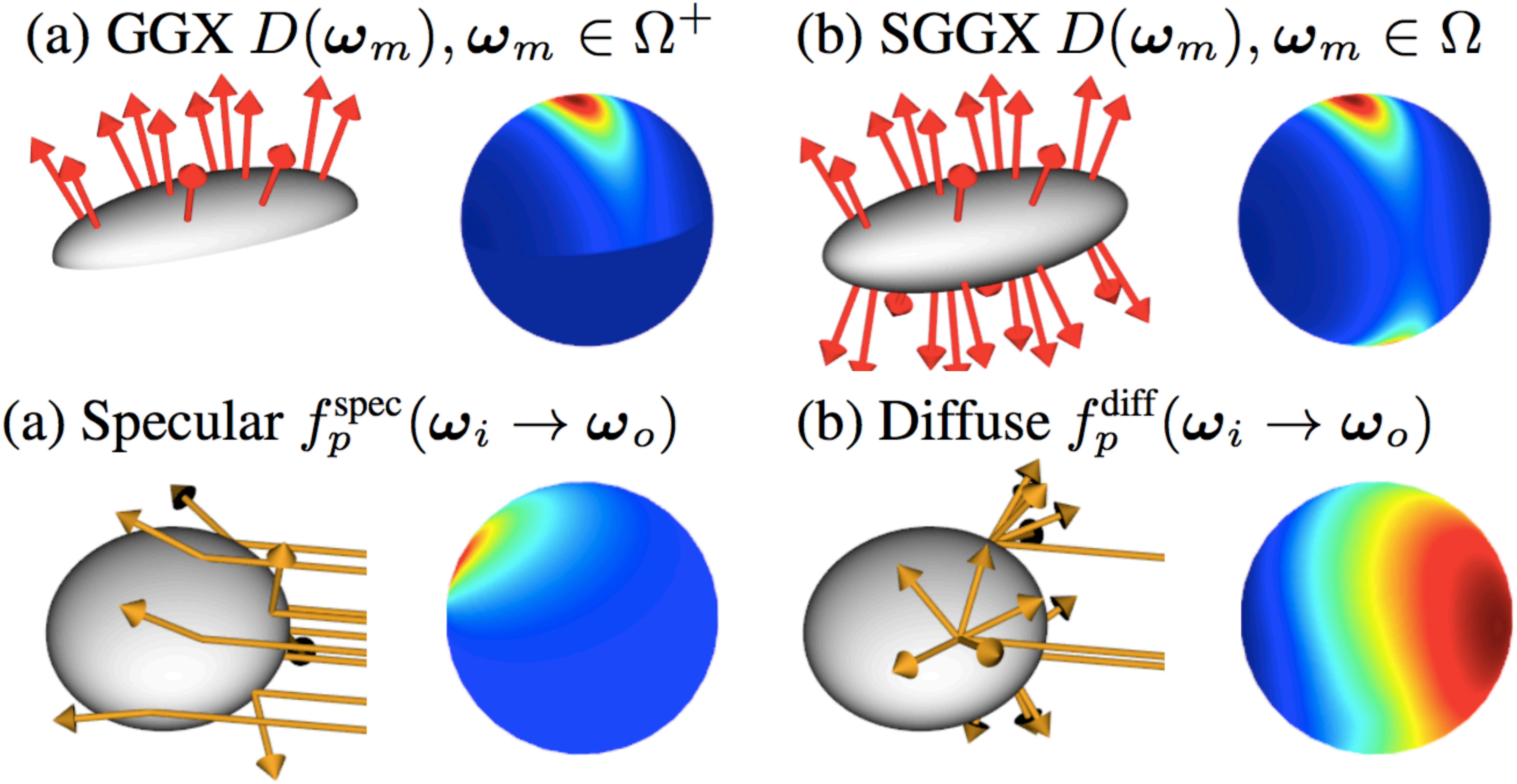
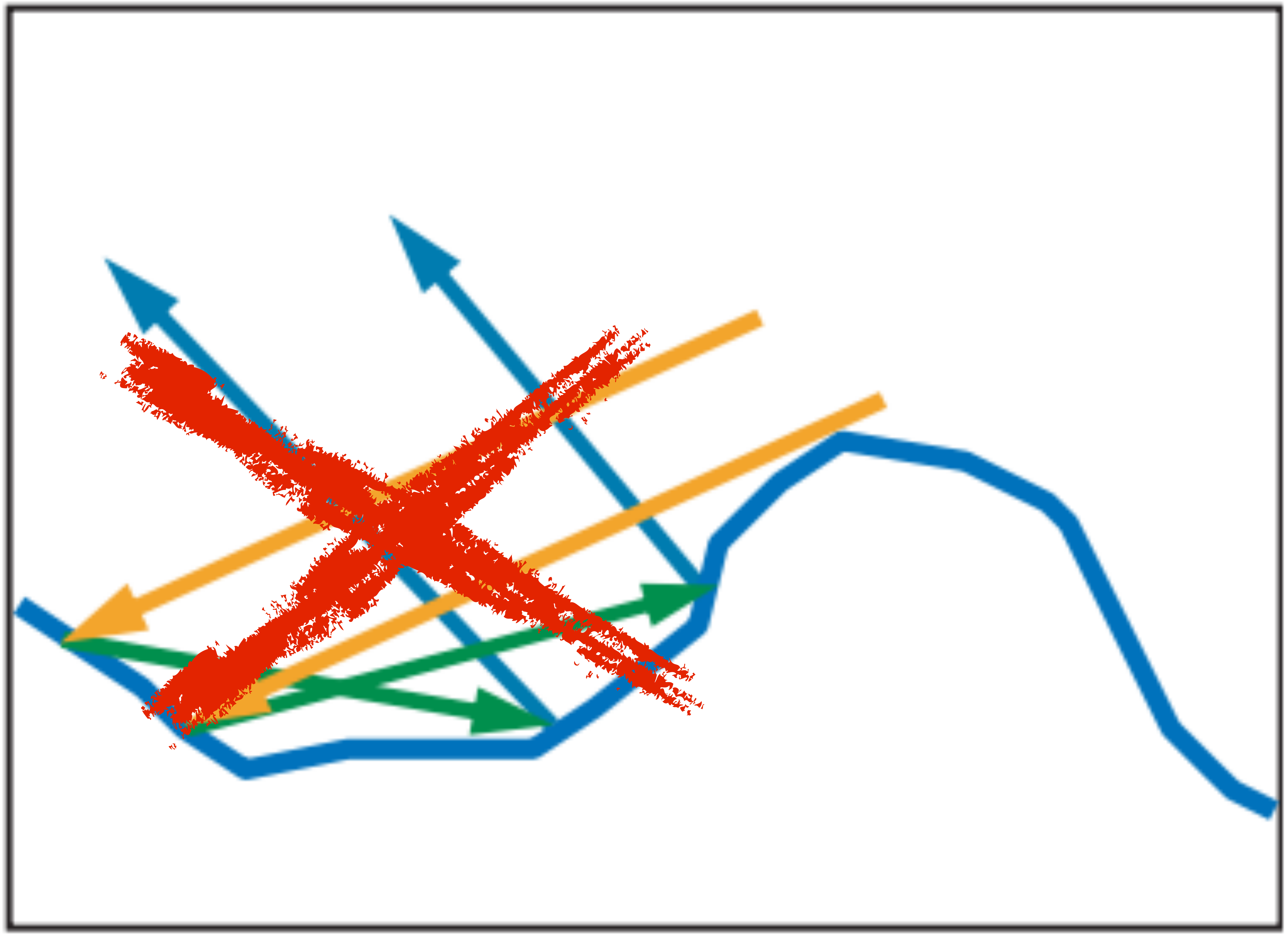


Image from “The SGGX Microflake Distribution”, Heitz et al., SIGGRAPH 2015

In 2015 Heitz et al. introduced the SGGX microflake distribution, an extension of the GGX microfacet distribution to the full sphere. It had many advantages over existing microflake distributions, allowing for linear interpolation, analytical evaluation and importance sampling of visible normals.

Multiple Surface Scattering



Render the Possibilities
SIGGRAPH2016

Now we'll switch gears for a moment. As I mentioned earlier, the lack of multiple bounce scattering is one of the fundamental limitations of microfacet theory.

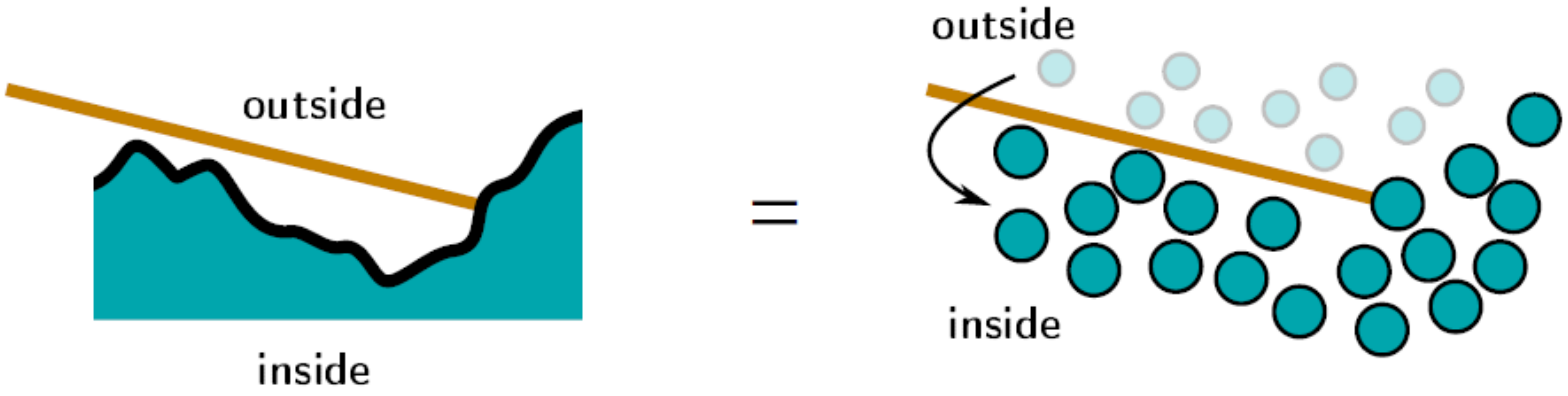
Multiple Surface Scattering: Analytical Models

$$f_{r,matte}(\vec{L},\vec{V}) = k(\lambda) \cdot \frac{(1 - a_{spec}(\theta')) \cdot (1 - a_{spec}(\theta))}{\pi(1 - a_{spec}^{ave})}$$

$$\rho = RE + R(1-E)RE + R^2(1-E)^2RE + \cdots = \sum_{i=0}^{\infty} R^k(1-E)^k = \frac{RE}{1 - R(1-E)}$$

There have been some attempts at corrective factors to add back the missing energy. The upper one is from the Eurographics 2001 paper by Kelemen & Szirmay-Kalos, and the lower one is the multiple scattering term for metals from the SIGGRAPH 2014 paper on layered materials by Jakob et al. (the paper also has a term for dielectrics).

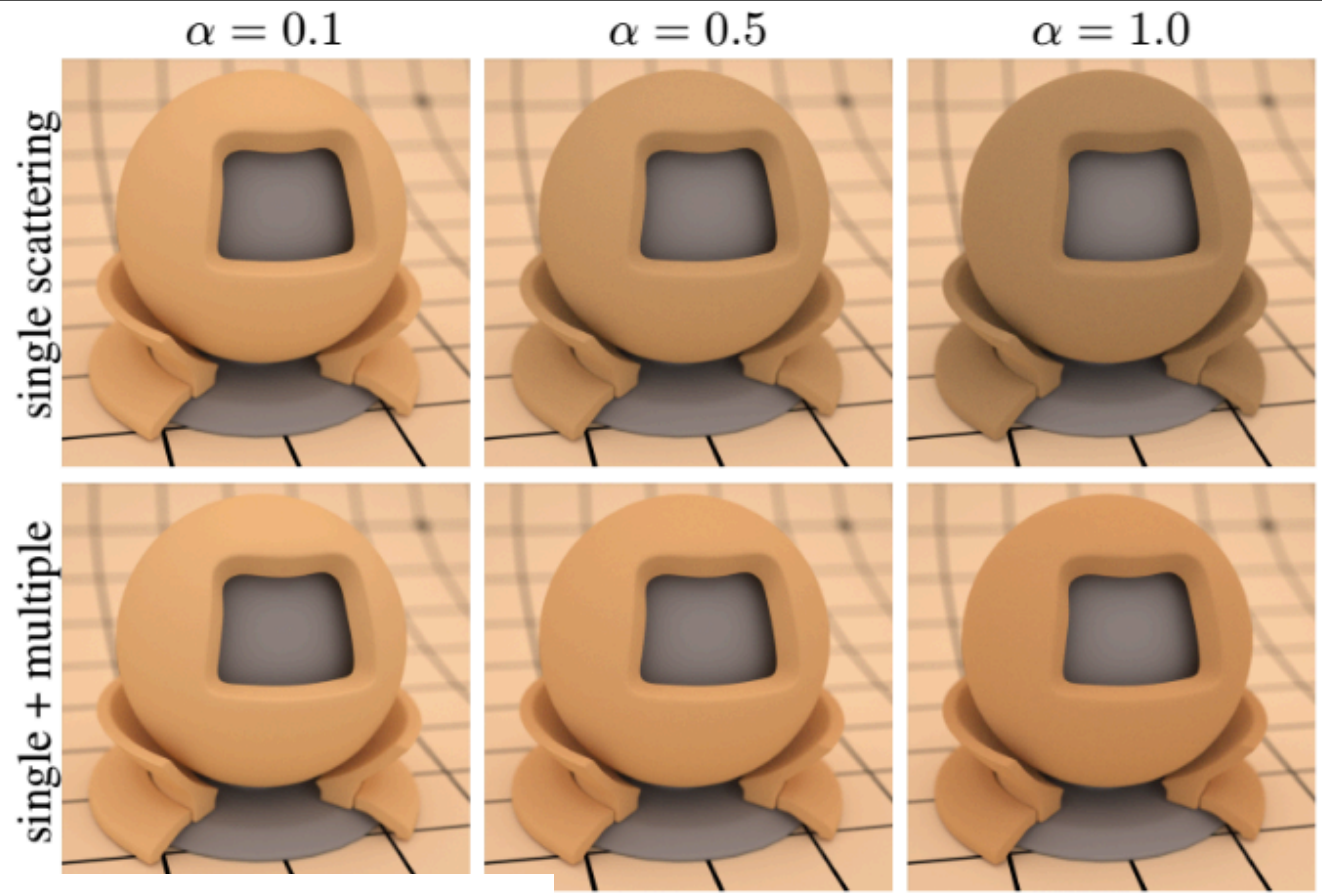
Multiple-Scattering Microfacet BSDFs with the Smith Model (SIGGRAPH 2016)

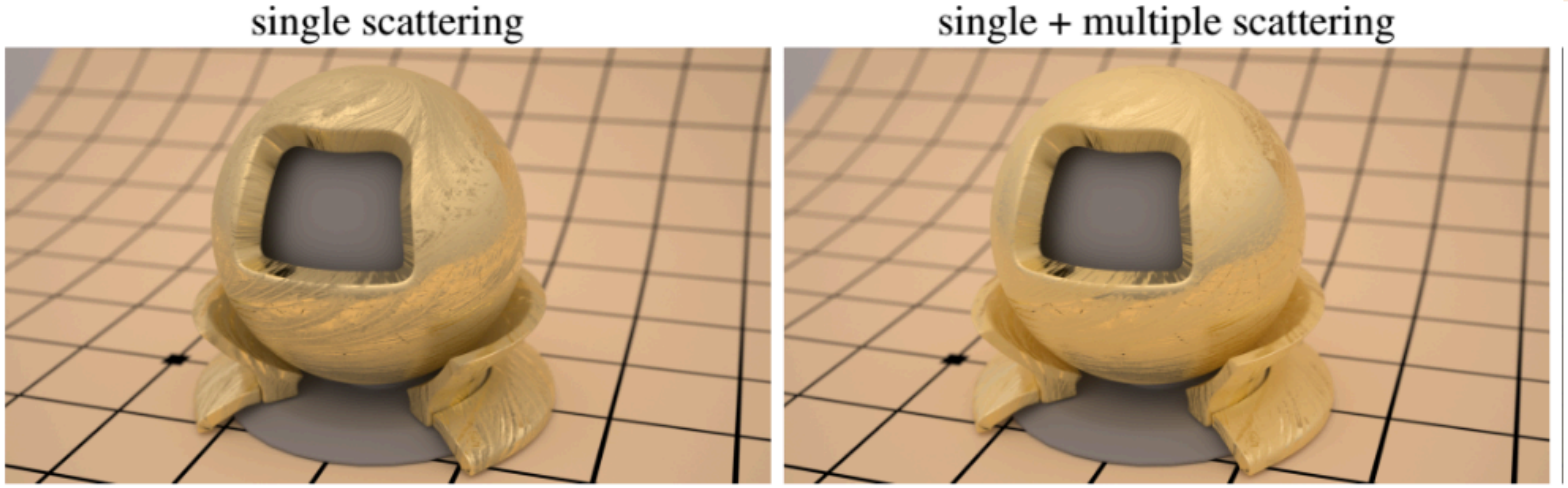


Images from “Multiple-Scattering Microfacet BSDFs with the Smith Model”, Heitz et al., SIGGRAPH 2016

And now tying together the two topics of microflakes and multiple surface scattering, this rather ingenious paper models a surface as a microflake volume with certain properties. Methods typically used to render participating media can be used to effectively render multiple surface bounces on a microfacet surface. Unfortunately, this model is stochastic so it’s not suitable for real-time rendering applications.

Images from “Multiple-Scattering Microfacet BSDFs with the Smith Model”, Heitz et al., SIGGRAPH 2016





Render the Possibilities SIGGRAPH 2016

The paper also showed what multiple scattering looks like — what we’re missing by not having it. Now rougher surfaces increase in saturation, instead of becoming darker due to lost energy. This is especially noticeable in the case of Spatially varying roughness.

Additional Progress Towards the Unification of Microfacet and Microflake Theories

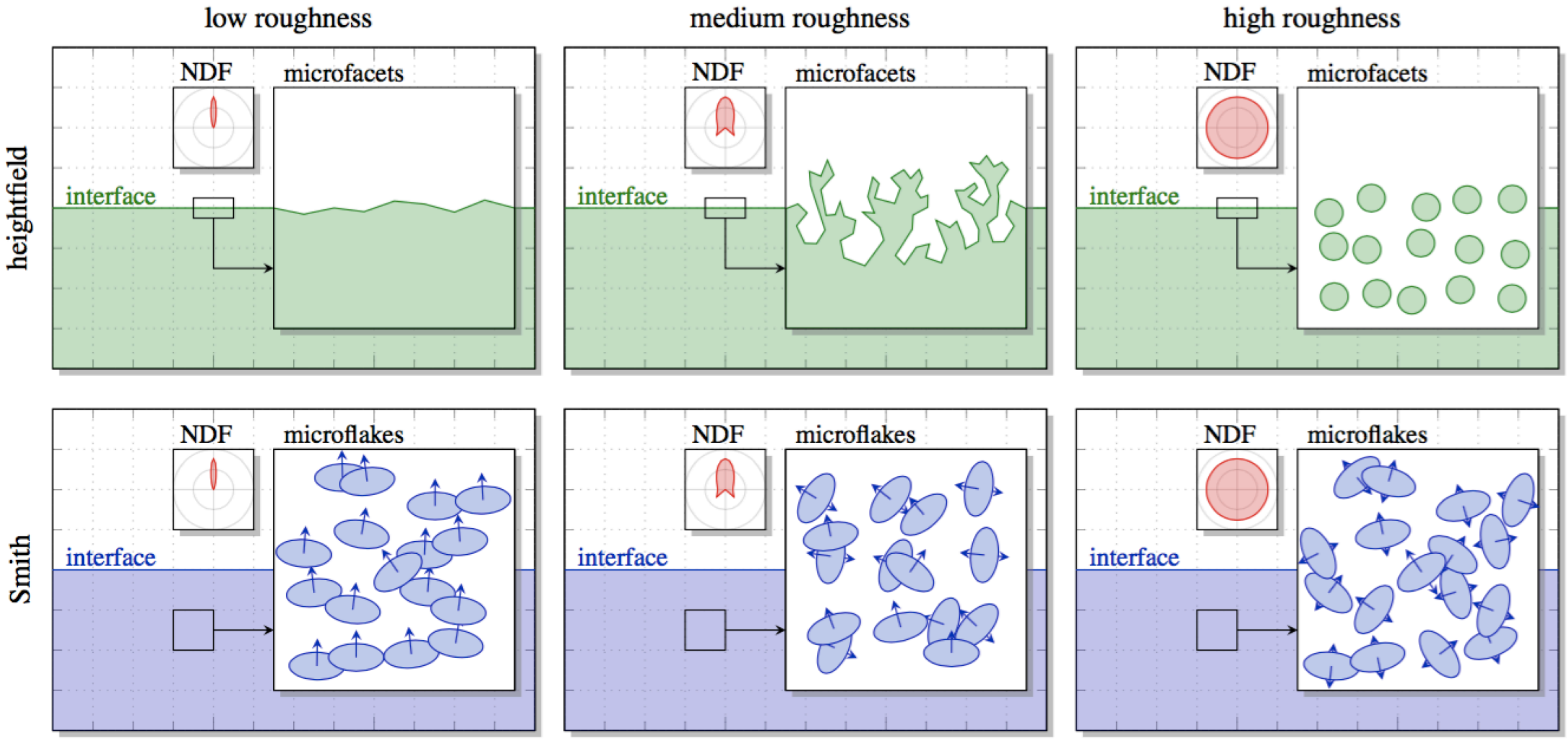


Image from “Additional Progress Towards the Unification of Microfacet and Microflake Theories”, Dupuy et al., EGSR 2016

This paper did a simpler derivation of surface multi-scattering, also based on microflakes but using a semi-infinite homogeneous volume instead of the variable-density volume used in the earlier paper. The authors also discuss potentially extending the NDFs beyond the hemisphere to model extremely rough or porous surfaces, and to create a continuum between surface and volume modeling, as shown in the figure.

Coarse Microgeometry

Render the Possibilities
SIGGRAPH 2016

The ability to model coarse or “glinty” microgeometry was another one of the limitations I discussed at the beginning of the talk. I’ll quickly go over some recent work that offers solutions in this area.

Discrete Stochastic Microfacet Models (SIGGRAPH 2014)

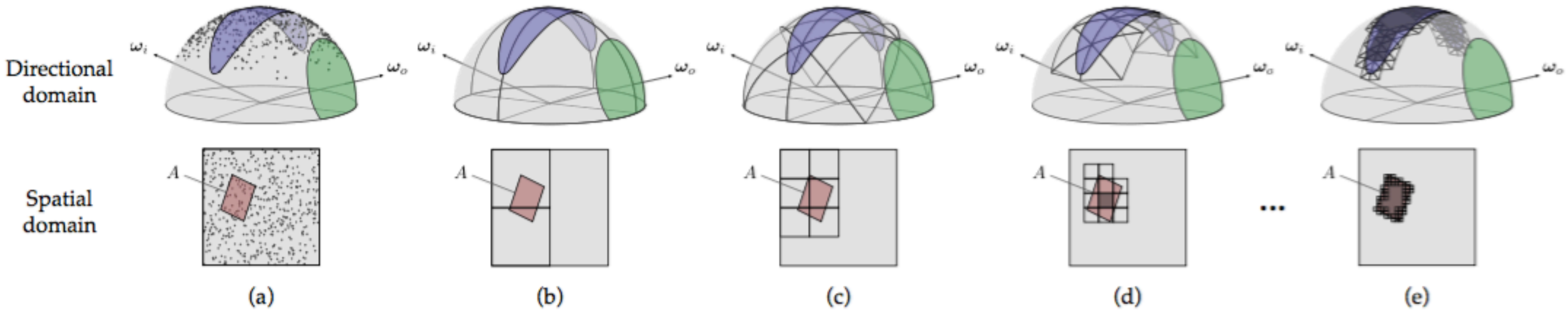


Image from “Discrete Stochastic Microfacet Models”, Jakob et al., SIGGRAPH 2014

Render the Possibilities
SIGGRAPH2016

This paper uses a microfacet model with a discrete distribution of scattering particles instead of the usual continuous distribution. The particles are generated stochastically in a temporally coherent way, over a 4D domain that includes the pixel footprint on the surface and the set of microfacet directions that reflect light into a small cone of outgoing directions.

Multiscale BRDF

An important concept introduced in this paper is that of a *multiscale BRDF*, defined over patches of surface and cones of area instead of infinitesimal points and rays. This concept is actually closer to the way BRDFs are implemented in practice, and given the intimate relationship of BRDF models to scale, I consider the multiscale BRDF to be in a way more fundamental than the BRDF as originally defined.

Real-time Rendering of Procedural Multiscale Materials (I3D 2016)

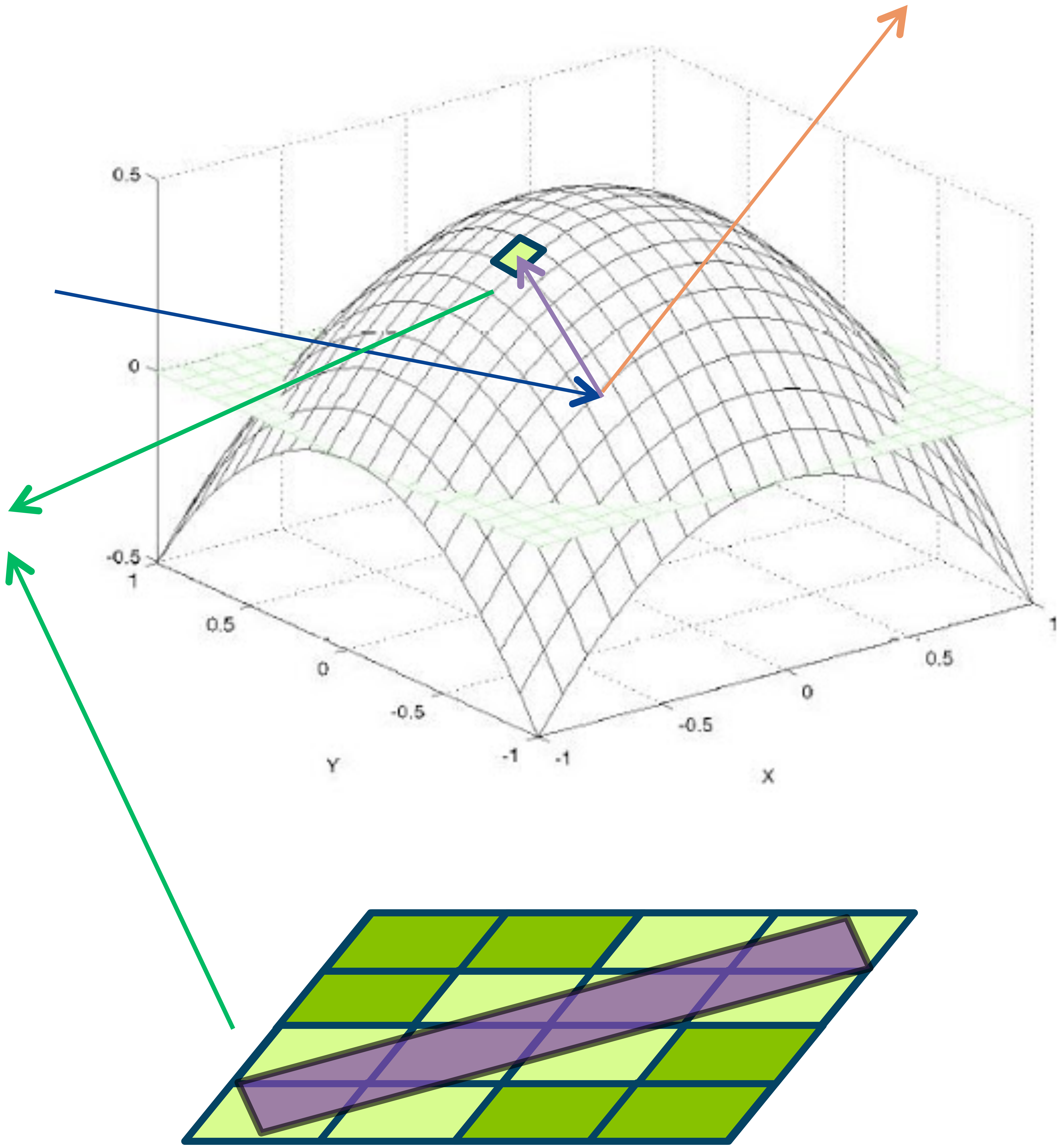
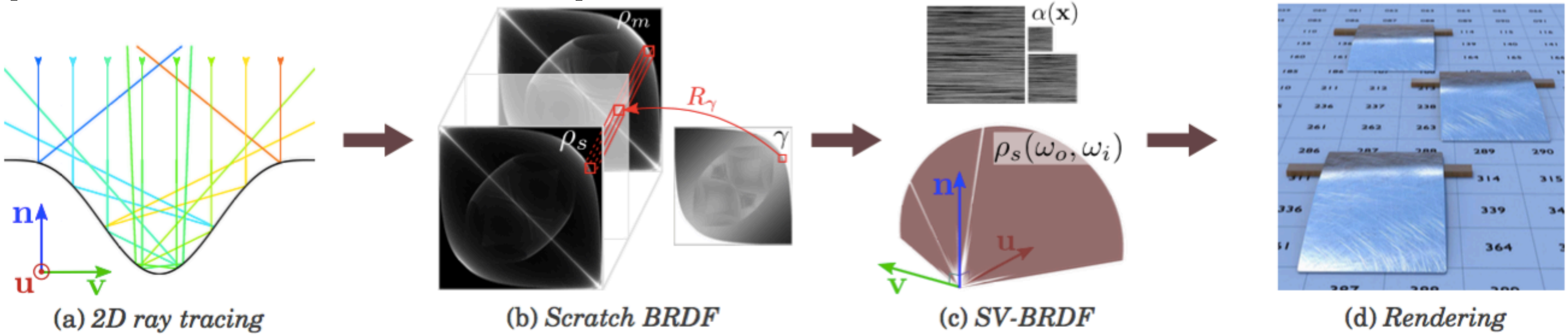
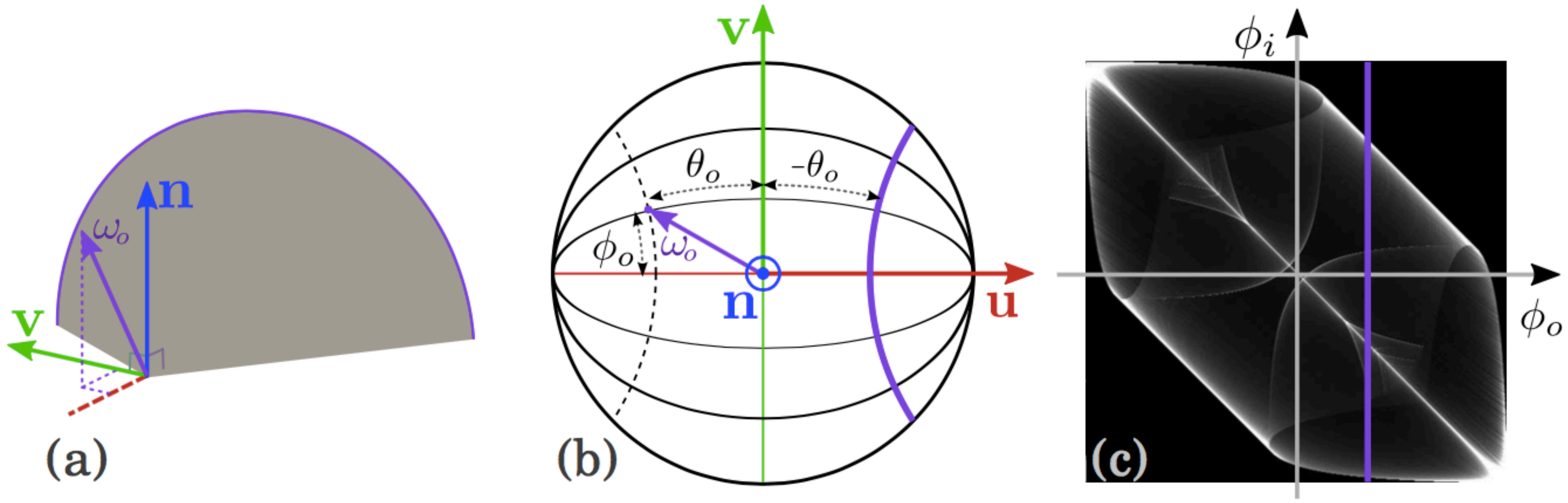


Image from “Real-time Rendering of Procedural Multiscale Materials”, Zirr & Kaplanyan, I3D 2016

The technique in the previous paper was too slow for real-time rendering applications. This paper speeds things up by carefully designing a real-time-friendly datastructure for evaluating a multiscale NDF.

Multi-Scale Rendering of Scratched Materials using a Structured SV-BRDF Model (SIGGRAPH 2016)





Images from "Multi-Scale Rendering of Scratched Materials using a Structured SV-BRDF Model", Raynond et al., SIGGRAPH 2016

Render the Possibilities
SIGGRAPH2016

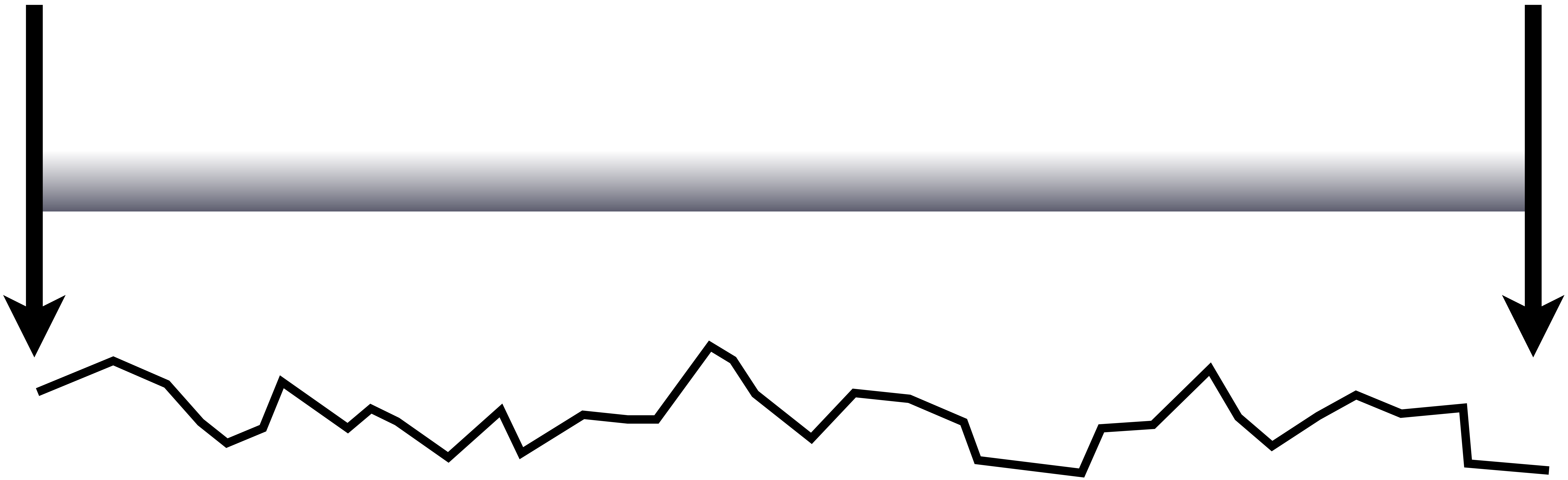
The technique in this paper is specialized to scratches specifically. In the top row we can see the steps of the algorithm. The authors precompute various scratch BRDFs (including multiple bounces within each scratch, which ties into a previous topic as well). This is stored with a novel 2D parameterization (shown on the bottom row), and then used in combination with a local scratch density to build a spatially varying BRDF which is used to render the surface.

Physical Optics (Wave) Models

Render the Possibilities
SIGGRAPH2016

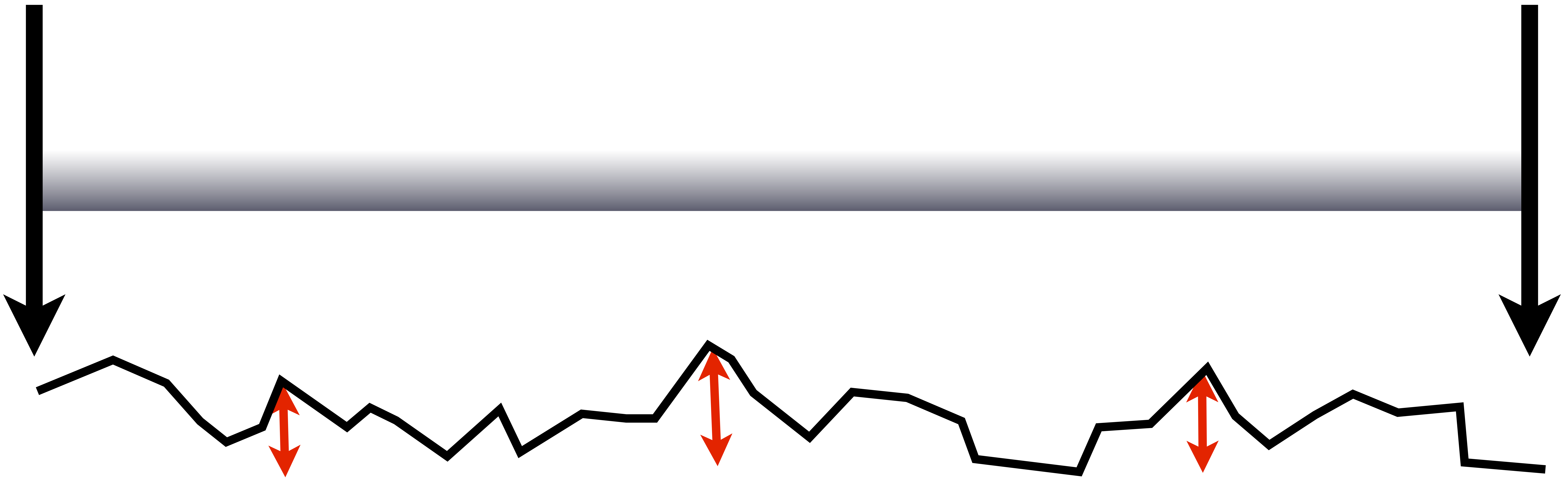
The last part of this talk will be about addressing the other fundamental limitation of microfacet theory, looking at ways to use wave optics to model reflectance. First I want to correct and extend some of the comments I made about wave reflectance last year; I now understand the topic a bit better (though there is still a lot to learn!)

Light Waves and Surface Scale



We'll start with light coming in at normal incidence, for simplicity. There are a few different domains of surface scale that have different effects on the light. But first it's important to note that not all scale is created equal. It turns out that scale in the direction of light propagation (surface height, in this case where light is coming in at normal incidence) is different than scale perpendicular to the light direction (in this case, scale along the surface).

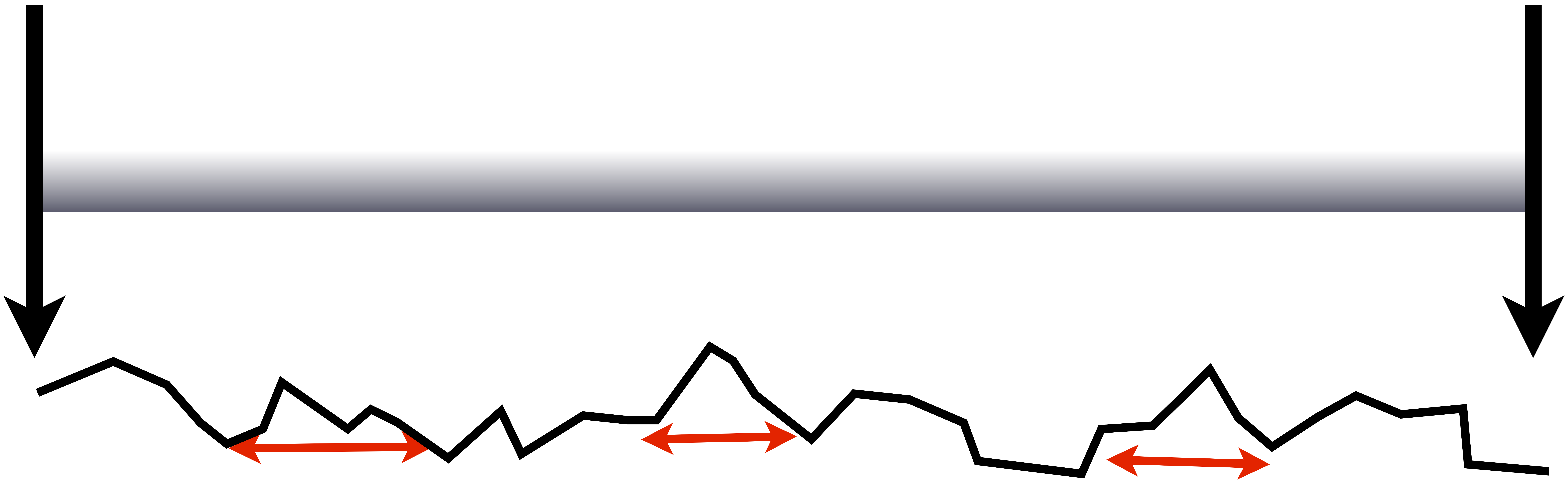
Light Waves and Surface Scale



Render the Possibilities
SIGGRAPH2016

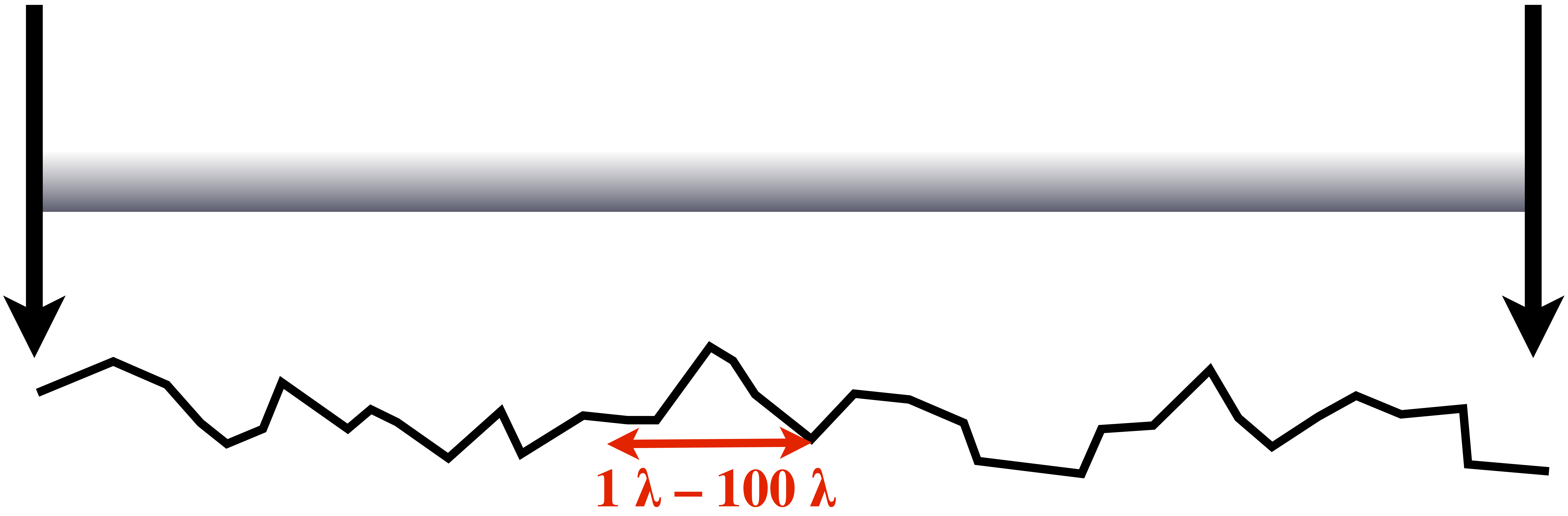
Height scale is just a matter of degree. More height causes a larger effect, less height causes a smaller effect — there are no cutoff points where we move into a new scale domain and different phenomena start happening.

Light Waves and Surface Scale



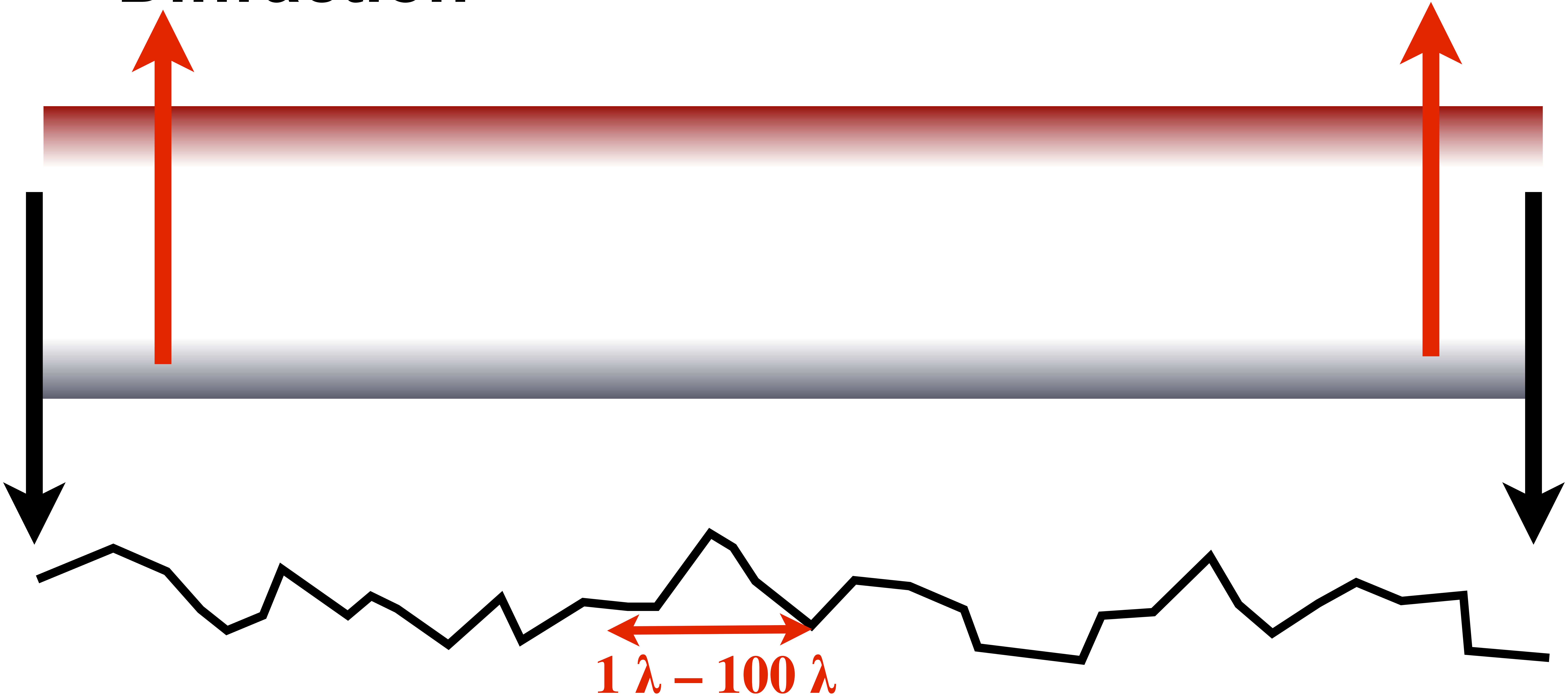
But when it comes to scale along the surface, that's a different matter. There definitely are different domains with different effects. But in one respect both kinds of scale are the same — they are both relative to the light wavelength. So what matters is not the absolute size of the surface detail, but how many multiples of the light wavelength it is.

Nanogeometry



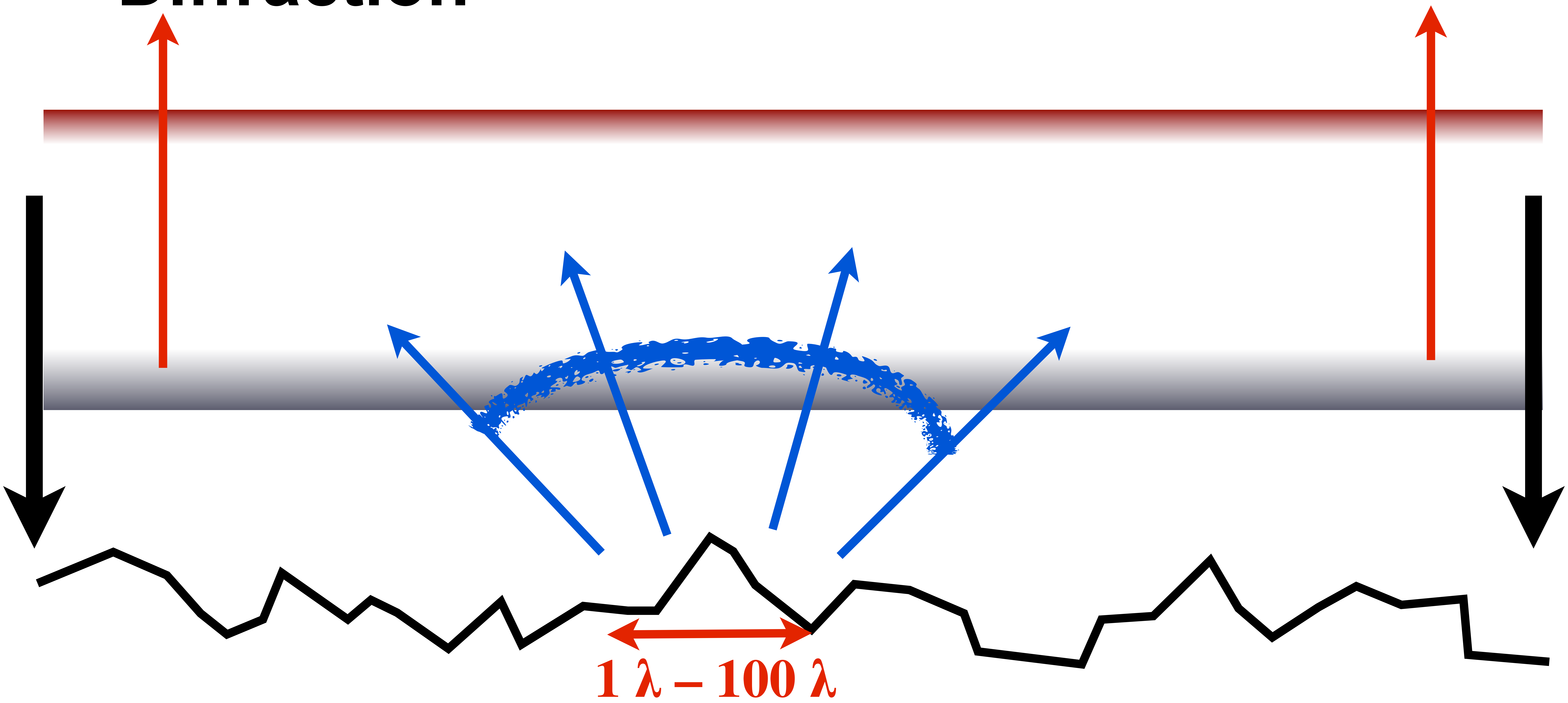
Since I haven't been able to find out any "proper" terms for these I will use completely non-standard terminology that I made up. I'll start with nanogeometry, a term I used last year to define geometry that causes diffraction. The relevant scale along the surface for nanogeometry is between one and about one hundred light wavelengths.

Diffraction



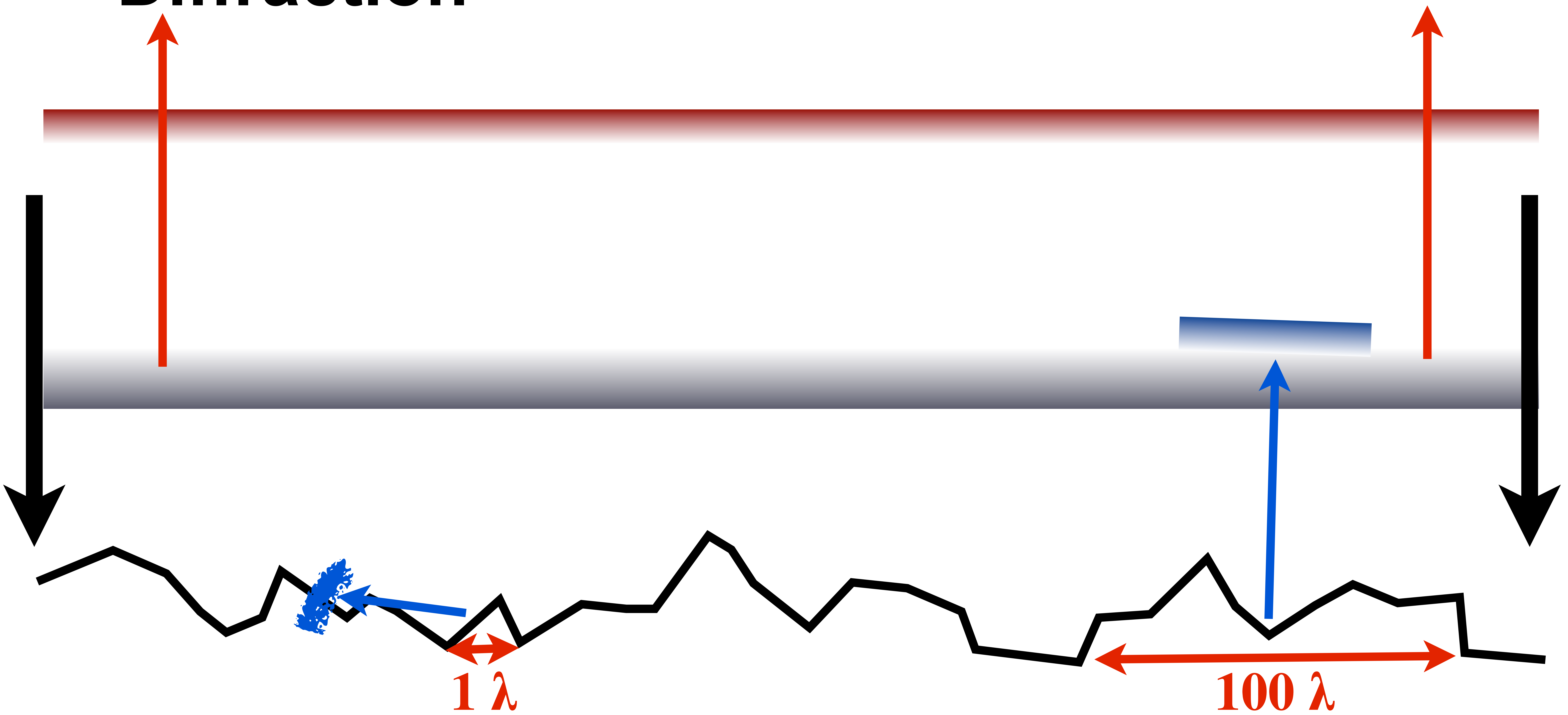
Very smooth surfaces like this one, with only this scale of roughness, return light back in the reflection direction like a perfect mirror. Diffraction doesn't spread this reflection...

Diffraction



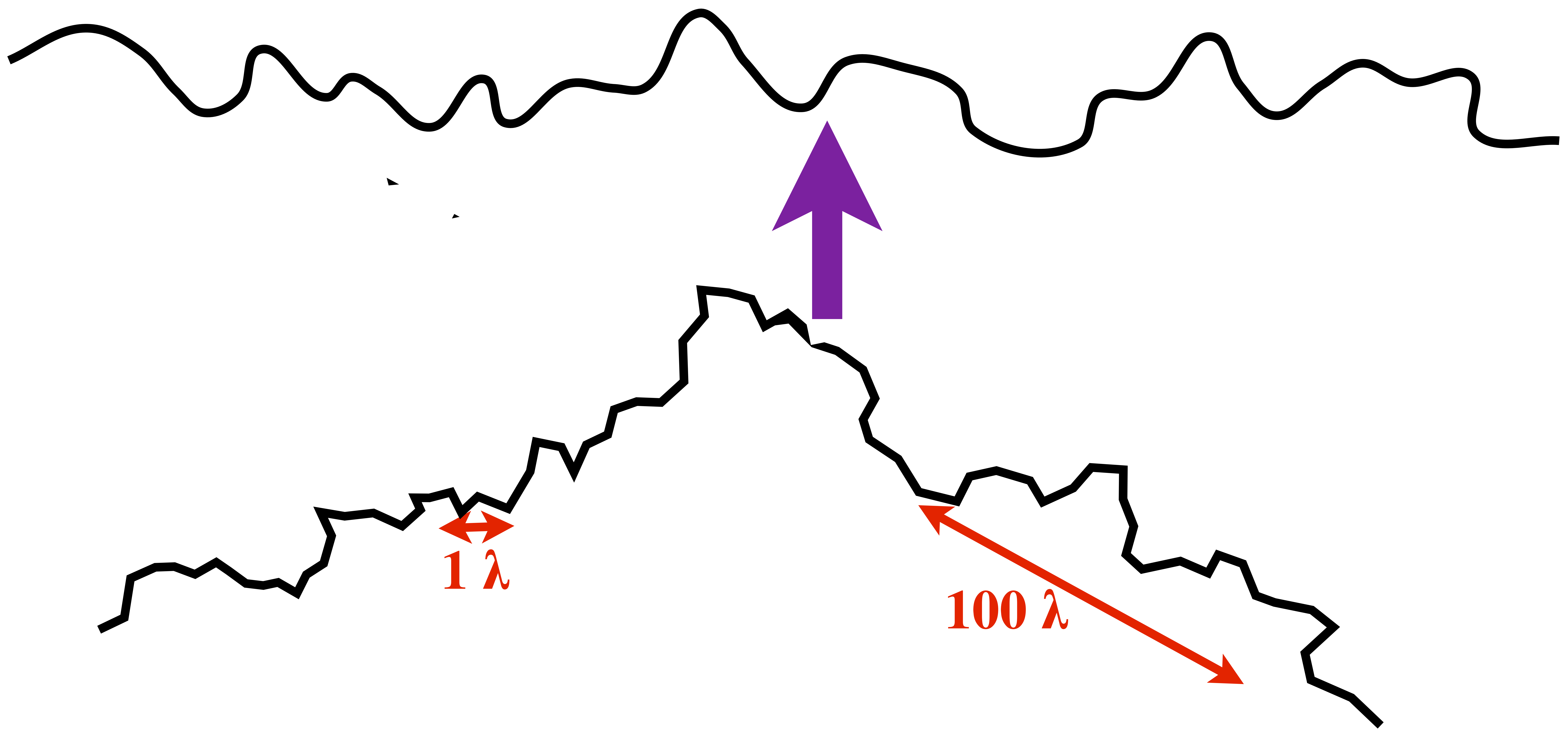
...it weakens it, stealing part of its energy and spreading that. So now you have a combination of a perfect delta-function spike of mirror reflection, and an incoherent cone of diffracted light. The amount of light diffracted depends on the height, and the angle of the diffracted light depends on the scale of surface roughness along the surface. Usually there's a mix of scales so the light is spread in a continuous cone. Since these angles depend on the scale relative to the light wavelength, that means red light and green light and yellow light etc. will all diffract at slightly different angles, which causes some angular color variation.

Diffraction



A bit counter-intuitively, the smaller the detail the wider the diffraction angle. The surface details that are 1 wavelength wide will diffract the light at 90 degrees; the ones that are around a hundred or more wavelengths wide will diffract the light a half degree or less so it starts merging indistinguishably into the mirror reflection. And that’s why scales outside that range have no effect. Light can’t be diffracted more than 90 degrees, so details smaller than a wavelength don’t diffract at all. And details much larger than 100 wavelengths don’t diffract at all either, for the opposite reason.

Band-Limiting



Since wavelengths outside this range don't affect diffraction, you can think of the diffraction as seeing the surface through a band-limiting filter, that throws away all the spatial frequencies that are too large, like the big bump in the middle of the bottom surface, or too small, like all the sharp jaggies in it. And the result is a surface like the top one. And the roughness height of <u>that</u> surface is the one that's relevant for calculating the amount of diffraction. Now all of this was for normal incidence - for other light angles just think of the relevant directions as shifting with the light, and otherwise the picture stays the same.

Band-Limiting

microgeometry

nanogeometry

picogeometry

Render the Possibilities
SIGGRAPH2016

Now the filtered detail doesn't go away entirely — the detail that's too big to be nanogeometry, too big to cause diffraction is the microgeometry we all know and love from microfacet theory. And the detail that's too small to cause diffraction I'll call picogeometry, which isn't an official term of any kind but that's what I've got. Picogeometry actually affects light in a different way, but that will have to wait for another day.

Why do we care?

So why do we care about this now, when we didn't for some decades? As I mentioned last year, there has been some work showing that everyday materials have some visible diffraction, and some of the "long tails" we are seeing are from that and not from the NDF.

A Physically-Based Reflectance Model Combining Reflection and Diffraction

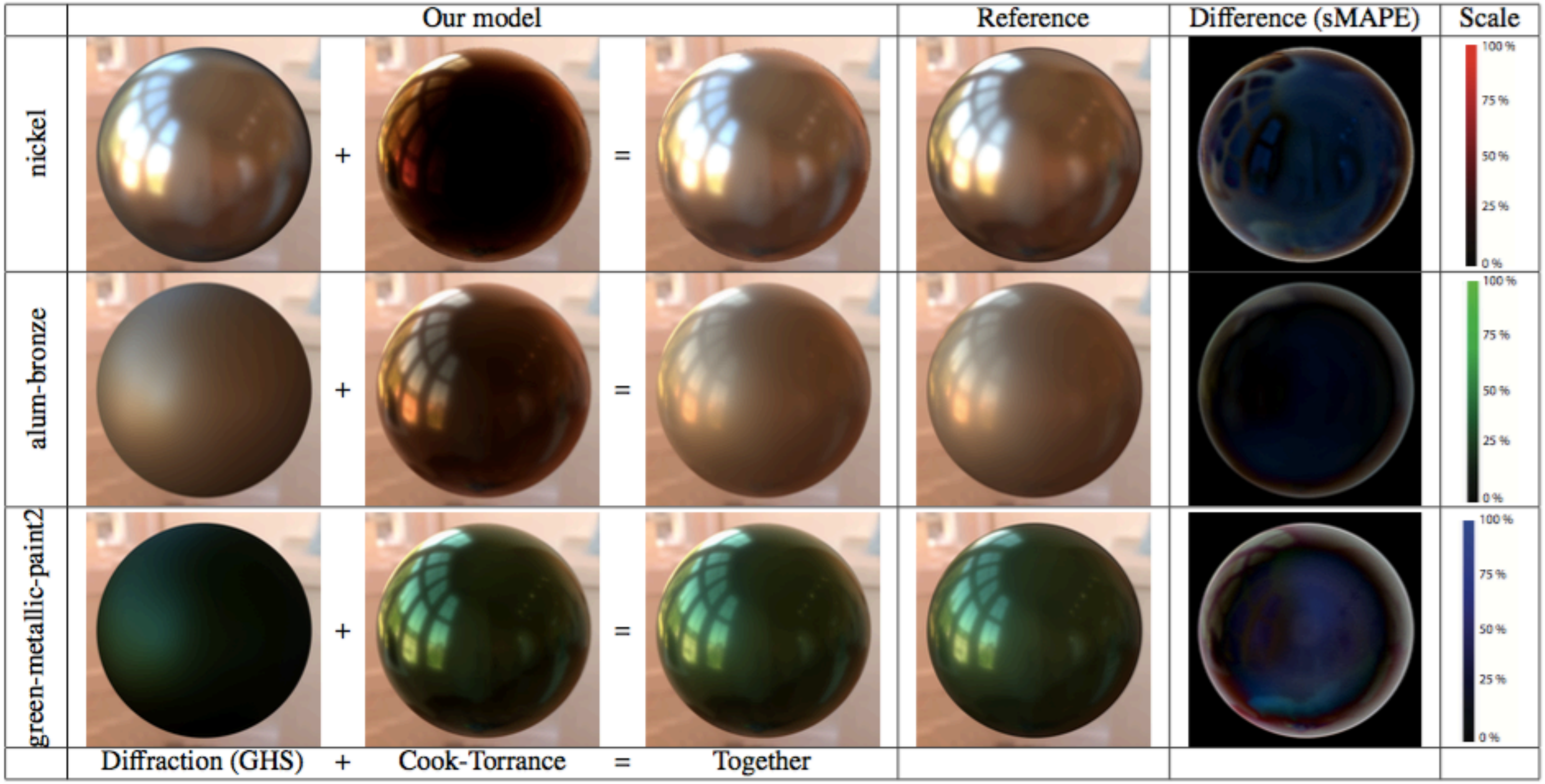


Image from “A Physically-Based Reflectance Model Combining Reflection and Diffraction”, Holzschuch & Pacanowski, INRIA Research Report 2016

This is a white paper by Nicolas Holzschuch & Romain Pacanowski which is about to be published in the next few days; Nicolas & Romain were kind enough to give me early access to it for the purpose of this talk. It’s where I got the Generalized Beckmann NDF I discussed earlier, but that’s a very minor part of the paper.

A Physically-Based Reflectance Model Combining Reflection and Diffraction

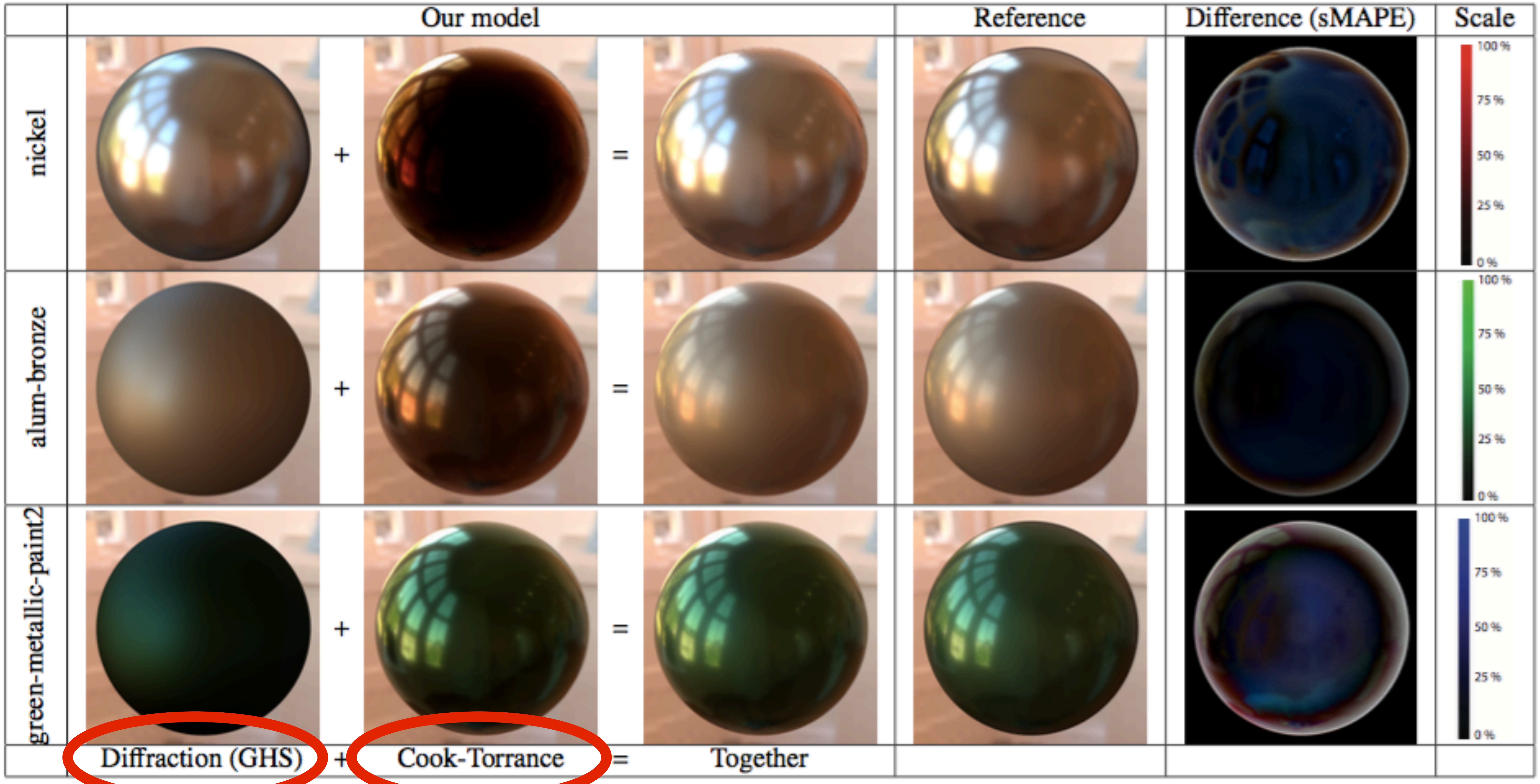


Image from “A Physically-Based Reflectance Model Combining Reflection and Diffraction”, Holzschuch & Pacanowski, INRIA Research Report 2016

This white paper has a model which combines a Cook-Torrance style microfacet model (this is the part that uses Generalized Beckmann) with a diffraction model using the latest optical theory - the Generalized Harvey-Shack (GHS) theory, dating from 2006. For context, most of the previous wave optics work in computer graphics has been using theory from the 60s, or the 80s. And GHS is pretty important — all the previous theories had limitations to only smooth surfaces, or only small incidence angles, etc. GHS is the first fully general theory of surface diffraction. And this paper combines it with microfacet theory by treating the GHS BRDF as the BRDF for the individual microfacets, which makes sense considering the band-limiting stuff I talked about earlier. We’re out of time so I won’t be able to go into more detail, but the white paper itself should be up really soon and I recommend you read it if you want to learn about this new direction in shading models.

Acknowledgements

- Steve Hill: assistance with course notes & slides, WebGL framework used for Fresnel visualization
- Brent Burley, Paul Edelstein, Yoshiharu Gotanda, Eric Heitz, Christophe Hery, Sébastien Lagarde, Dimitar Lazarov, Cedric Perthuis, Brian Smits: inspirational discussions on physically based shading models
- Nicolas Holzschuch and Romain Pacanowski: for early access to, and illuminating discussion about, their latest work on combining physical and geometric optics models

To wrap up, I'd like to thank some people who helped me with this talk, and thank you for listening.