Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2011-052 December 14, 2011

A Frequency Analysis of Monte-Carlo and other Numerical Integration Schemes

Fredo Durand

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

A Frequency Analysis of Monte-Carlo and other Numerical Integration Schemes

Frédo Durand
MIT CSAIL

Abstract

The numerical calculation of integrals is central to many computer graphics algorithms such as Monte-Carlo Ray Tracing. We show that such methods can be studied using Fourier analysis. Numerical error is shown to correspond to aliasing and the link between properties of the sampling pattern and the integrand is studied. The approach also permits the unified study of image aliasing and numerical integration, by considering a multidimensional domain where some dimensions are integrated while others are sampled.

Keywords: Numerical Analysis, Integration, Fourier, Monte-Carlo, Aliasing, Rendering, Ray Tracing

1 Introduction

Sampling is at the heart of two central issues in computer graphics: image antialiasing and Monte-Carlo integration. They are intimately related, and it has been clear that antialiasing involves some numerical integration since we usually can only point-sample the image function. However, to a few exception, there has been little connection between aliasing and Monte-Carlo integration.

In this text, we argue that Fourier analysis is a powerful tool to understand numerical integration techniques such as Monte-Carlo estimators. Fourier analysis can characterize both the integrand and the sampling pattern, as well as their interaction. Fourier analysis is also the tool of choice to study aliasing, a critical issue in computer graphics where numerical integration is needed for each pixel to generate a discrete 2D images that suffers not only from integration noise at each pixel, but also from 2D aliasing across pixels.

The present text focuses on theoretical insights. It has been work in progress for way too long, and I have decided to publish the current state as a tech report. The list of open issues is still long and, in particular, includes importance sampling and the convergence of estimators when the sampling rate is increased.

We hope that the combination of the present perspective on numerical integration and recent frequency analysis of light transport and other rendering effects, e.g. [Ramamoorthi and Hanrahan 2001; Durand et al. 2005; Egan et al. 2009; Soler et al. 2009; Egan et al. 2011a; Egan et al. 2011b]

The insights in this document have been used to derive some of the sampling rate in work on the Fourier analysis of motion blur [Egan et al. 2009]. The careful reader will have noticed a factor of 2 in the sampling rate we derived there, which comes from the revised Shannon sampling theorem derived in the present document.

1.1 Overview

We express Monte-Carlo and numerical integration in terms of signal processing, where a sampling step is followed by a summation step. Sampling potentially results in aliasing, which results in error. Precisely, we show that numerical error is the aliasing at the DC. Said differently, it is the sum of the frequency-by-frequency product of the integrand spectrum and the sampling pattern spectrum (the dot product between the two spectra). A final expression is that error corresponds to the correlation between the integrand and the sampling pattern.

This allows us to study how different types of integrands and sampling patterns interact. In particular, we derive a revised sampling theorem that states that a function can be perfectly integrated if it is sampled at at least its maximum frequency. This is a factor of two compared to Nyquist and Shannon.

We then study the multidimensional case involved in image synthesis, where each pixel is both a sample of a continuous function and the result of a numerical integration across dimensions such as time (for motion blur), lens (depth of field), or incoming illumination. We show that the Fourier analysis of the higher-dimensional case (2D xy times time, times lens, etc.) allows us to express both antialiasing issues and numerical integration error. In particular, it shows how error due to numerical integration might result in structured artifacts when sampling is poorly chosen, such as the use of the same random sequence for all pixels.

1.2 Related Work

The two most related papers are Cook's stochastic sampling [1986] where he studies the effect of non-uniform sampling and compares blue noise and jittering, and Mitchell's spectrally optimal sampling [1991]. In particular, Mitchell's work studies numerical integration using Fourier analysis, but we feel that the signal processing perspective is sometimes hidden below the surface and we seek to emphasize it in this document.

Ouellette and Fiume [2001] look at numerical integration and compare various estimators for 1D integrals, specifically in the context of linear light source. They characterize the spectrum of the sampling patterns but do not use it directly for error analysis.

Fourier analysis has been applied to the numerical computation of integrals of periodic functions, [Gautschi 1997] p155 [Boyd 2011] p457. Usually restricted to the trapezoid rule and not a full signal processing/aliasing perspective. It has also been shown that the numerical error predicted by various schemes such as trapezoid or Simpson can be extremely conservative for periodic function, and that considering functions in terms of their Fourier series can be used to derive better error bounds for periodic functions [Weideman 2002].

2 Basic Monte-Carlo Integration (1D, uniform sampling)

We first study the simple case of a 1D integral and show that it can be expressed in terms of signal processing and that error corresponds to aliasing.

We seek to evaluate

$$I = \int_0^1 f(x)\,\mathrm{d}x \tag{1}$$

This is, by definition, the DC of the Fourier transform of f.

$$I = \hat{f}(0) \tag{2}$$

A numerical scheme such as a Monte-Carlo estimator with N uniform samples is:

$$I_N = \frac{1}{N}\sum f(x_i) \tag{3}$$

where the x_i are random samples distributed uniformly across the domain. In this document, we focus on one instance of this integrator (for a given set of N samples) while the Monte-Carlo literature usually considers the random *process* where the x_i are random variables. In our case, the x_i can come from a random sequence generator, but might also correspond to other numerical schemes such as trapezoids or Simpson.

A more general integrator has weights for each samples, which can come from importance sampling or other derivations such as polynomial approximations in the case of Simpson.

$$I_N = \sum w_i f(x_i) \tag{4}$$

The weights often sum to 1, but not always. For example, importance sampling exists both in a normalized or unnormalized form.

2.1 Sampling theory

We separate the numerical scheme into two steps, sampling and summation. Similar to sampling theory, we define the sampling function

$$S(x) = \sum w_i \delta(x - x_i) \tag{5}$$

where δ is the Dirac characteristic function. That is, S is non-zero only at the x_i and integrates to 1 if the w_i sum to 1. If the x_i are regularly distributed, we get the traditional Dirac train used in uniform sampling. For Monte-Carlo integration, the x_i are randomly distributed and S is Poisson noise.

Estimators are usually expressed as a discrete sum over the x_i. However, now that we have sampled the integrand, we can consider that the estimator is a continuous integral:

$$I_N = \int S(x) f(x)\,\mathrm{d}x \tag{6}$$

That is, we first *sample* the function f and we then take the integral, which corresponds to the DC of the sampled integrand.

We then express this succession of sampling and integration in the Fourier domain. Sampling is a multiplication in the primal, and therefore a convolution in the Fourier domain

$$\widehat{S.f} = \hat{S} \otimes \hat{f} \tag{7}$$

This convolution generates well-known aliasing effects. If S is the regular sampling impulse train, then $\hat{S}$ is also a regular impulse train and we obtain the traditional replicas of $\hat{f}$ in the Fourier domain. If S is a Poisson distribution of samples, its Fourier spectrum has energy at all frequencies, and the effect of aliasing is more "diffuse."

We can now express Monte-Carlo integration or any other sampling-based numerical scheme by noting that Equation 6 is the DC of the sampled integrand:

$$I_N = \left(\hat{S} \otimes \hat{f}\right)(0) \tag{8}$$

In summary:

Monte-Carlo integration can be seen in the Fourier domain as a convolution by the spectrum of the sampling pattern followed by the extraction of the DC (value at frequency 0).

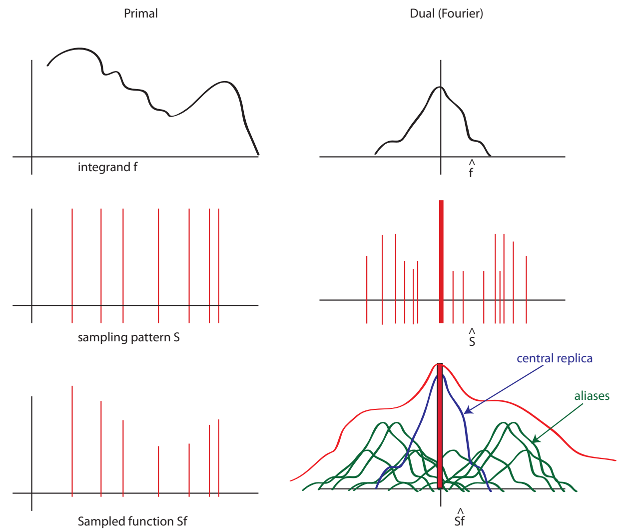

Figure 1: *Fourier pipeline for Monte-Carlo integration.*

Dot product perspective We expand the convolution integral and express equation 8 as:

$$I_N = \int \hat{S}(\omega)\hat{f}(-\omega)\,\mathrm{d}\omega \tag{9}$$

and because f is real we can replace $\hat{f}(-\omega)$ by the conjugate of $\hat{f}$:

$$I_N = \int \hat{S}(\omega)\hat{f}^*(\omega)\,\mathrm{d}\omega \tag{10}$$

The numerical estimation of the integral of f using a sampling pattern S is the dot product between their spectra.

This is not surprising since it is also a dot product in the primal, and the Fourier transform is simply an orthonormal change of basis, which preserve dot products.

As an extreme case of this formula, consider the original integral, which can be expressed as the sampling by a constant function. That is, the sampled function is the integrand multiplied by $S(x) = 1$. In this case, the Fourier transform of S is a Dirac and is zero everywhere else, leading to an error of 0.

2.2 Error analysis

We study numerical error due to a given sampling pattern S. Compared to the true integral in Eq. 2, the only difference is the convolution by the sampling pattern. We first consider the common case where the spectrum of S has a DC component $\hat{S}(0) = 1$, because the w_i some to 1. This means that the error in Monte-Carlo integration is due to the other parts of the spectrum of S.

The error in Monte-Carlo integration can be expressed in the Fourier domain as the aliasing caused by the sampling pattern at the DC.

$$I_N - I = \hat{f}(0) - \left(\hat{S} \otimes \hat{f}\right)(0) \tag{11}$$

In the dot product perspective, we first consider sampling patterns that sum to one. In this case, the DC term ($\omega = 0$) in equation 10 is the exact integral. This means that **error is the dot product between the spectrum of f and the spectrum of S where the DC is replaced by zero.**

Expressed differently, **error is the correlation between the sampling pattern and the integrand spectra.**

If the sampling pattern does not sum to one, which can happen with importance sampling, **the error has an additional term, which is the true integral multiplied by the difference between the DC of the sampling pattern and one.**

2.3 Regular sampling (Trapezoid): revised Nyquist criterion

When the x_i are regularly distributed over the domain, we obtain Trapezoid integration[1]. The spectrum $\hat{S}$ is also a Dirac comb, and we are in the presence of traditional aliasing where the original spectrum is regularly replicated. This means that the integration error is due to the frequency content at the sampling rate and its multiples. For example, if the function to be integrated is a sine wave at the sampling frequency, we get a systematic bias because $f(x_i)$ is the same value for all x_i.

However that if the integrand is band-limited and the sampling rate of our estimator satisfies the Nyquist criterion, the integration is exact. Not surprising, but always good to know.

Better still if we are under *twice* the Nyquist limit, the integral is still correct because aliasing occurs but does not affect the DC.

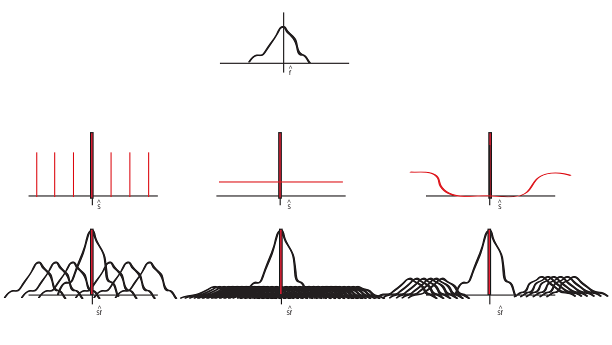

Figure 2: *The aliasing due to different sampling patterns.*

An integral can be numerically estimated exactly if the integrand is regularly sampled at a rate that is higher than its maximum frequency (half the Nyquist rate).

A similar result is known in the numerical integration literature where the integration of trigonometric polynomials is studied, e.g. [Gautschi 1997] p. 155, but is usually not expressed in terms of sampling rate and aliasing in signal processing.

2.4 Poisson sampling and Monte-Carlo integration

Poisson sampling In basic Monte-Carlo integration, the x_i are purely random and independent. This means that S is Poisson and its spectrum $\hat{S}$ is flat (except at the DC) and the phase $\phi(\omega)$ is random.

This sampling pattern is the most agnostic of all: it creates aliasing or all frequencies, all frequencies equally contribute to error. As a result, it's very general and makes no assumption on the frequency content of the integrand. However, if the integrand has less energy in the high frequencies, it can be wasteful.

Review of MC convergence and variance By construction, the contribution $f(x_i)$ of a sample of a Monte-Carlo estimator has the variance of the integrand f. Recall that when multiplying a function by a scalar, the variance gets multiplied by the square of the scalar: $\sigma^2(kf) = k^2\sigma^2(f)$. When adding two functions, their variances are added together as well as their covariance: $\sigma^2(f+g) = \sigma^2(f) + \sigma^2(g) + cov(f,g)$. In basic Monte-Carlo integration, the x_i are iid, which means that the their covariance and that of the $f(x_i)$ is zero. As a result, since we add N terms and multiply by $1/N$, we obtain

$$\sigma^2\left(\frac{1}{N}\sum f(x_i)\right) = \frac{1}{N}\sigma^2(f) \tag{12}$$

Since the error (standard deviation) is the square root of the variance, we obtain the $1/\sqrt{N}$ convergence of Monte-Carlo integration.

Variance and power spectrum Recall that the variance of a function is the integral of its power spectrum minus the DC:

$$\sigma^2(f) = E(f^2) - E(f)^2 \tag{13}$$

$$= \int f^2(x)\,\mathrm{d}x - DC(f)^2 \tag{14}$$

and according to Parseval's theorem, we have $\int f^2(x)\,\mathrm{d}x = \int \hat{f}^2(\omega)\,\mathrm{d}\omega$, therefore

$$\sigma^2(f) = \int |\hat{f}|^2(\omega)\,\mathrm{d}\omega - DC(f) \tag{15}$$

That is, the variance is the integral of the power spectrum except at the DC. This is already providing an insight about the link between Fourier analysis and Monte-Carlo integration.

Fourier analysis of MC convergence We have seen that the error is the dot product between the sampling spectrum and the integrand spectrum (minus the DC). This means that the squared error is

$$E[(I-I_N)^2] = E\left[\left(\int \hat{S}(\omega)\hat{f}(\omega)\,\mathrm{d}\omega\right)^2\right] \tag{16}$$

We expand the square and the expected value integral note that all the cross term vanish because $\hat{S}$ is random[2]. This gives

$$E[(I-I_N)^2] = \int E\left[\hat{S}^2(\omega)\hat{f}^2(-\omega)\right]\mathrm{d}\omega \tag{17}$$

and we can take the $\hat{f}$ terms out because the expected value is with respect to S:

$$E[(I-I_N)^2] = \int \hat{f}^2(-\omega)E\left[\hat{S}^2(\omega)\right]\mathrm{d}\omega \tag{18}$$

Since the expected power spectrum of S is flat (random samples), we obtain the variance multiplied by a constant that represents the power spectrum of S. Below we derive that this constant is $1/N$.

2.4.1 Simpler derivation

The power spectrum of a Poisson process is known to be a dirac $\nu^2\delta$ in the center and a ν flat spectrum for a Poisson process at rate ν.

[1] Regular sampling is also what you obtain with Quasi-Monte-Carlo integration with the Halton sequence for a power of two number of samples.

[2] I probably miss the $-\omega$ terms, which probably result in real value.

Worst case analysis when all phases interfere negatively with integrand spectrum and you get the sum of the absolute values of $\hat{f}$ and the L1 norm of the spectrum of the integrand (minus the DC as usual).

2.5 Study of different sampling patterns

Non-uniform weights For a fixed DC, the energy (variance) of a sampling pattern is bounded from below by the energy of the patterns where all weights are equal. Therefore the power spectrum of all sampling patterns is higher than that of the one with uniform weights (Parseval). This suggest that, without additional knowledge, the best strategy is to have all samples have the same weight.

Simpson The Simpson estimator is an extension of trapezoid that integrates second-order polynomials exactly. However, we show that its frequency properties are not ideal.

The *compound* Simpson scheme is:

$$I_N^s = \frac{1}{3N}\left(f(0) + f(1) + \sum 4f[(2i-1)h] + 2f(2ih)\right) + O(h^4) \quad (19)$$

This corresponds to a sampling pattern that has different weights for the various Diracs.

The Fourier transform is just a shifted version of that of the regular Dirac comb. This is in line with experiments that showed that Trapezoidal methods often do better than Simpson's scheme [Cruz-Uribe and Neugebauer 2002].

Blue Noise A blue noise sampling pattern has no low frequency (see Figure). As a results, the aliasing is restricted to the high frequencies. This pattern is great when the integrand has less energy in the high frequencies.

2.6 Error as a function of the integrand

Recall that the error of an integrator is the integral of the frequency-by-frequency product of the spectrum of the integrand and that of the sampling patterns. This means that if we know that some bands of frequency have more energy in the integrand, the sampling pattern should seek to have less energy there.

Natural signals are known to have a $1/\omega$ spectrum. The lower energy in the high frequencies means that aliasing there won't cause as much numerical error for integrators. This means that a sampling pattern that exhibits blue-noise properties

Similarly, uniform sampling improves better than $1/\sqrt{N}$ when the integrand has a falloff in the spectrum because it has no energy before the period of the sampling pattern.

3 Multi-dimensional case and dependent integrals

In computer graphics, Monte-Carlo estimators are usually not used to compute a single integral but to compute the values of a 2D array of pixels. The accuracy of each integral is not the only issue, and the visual noise introduced is paramount for image quality. For example, it is well known that using the same random sequence for all pixels in an image results in structured error that is extremely objectionable. How can we study the effect of the sampling strategies on image quality? We show below that it can be easily studied using multidimensional Fourier analysis.

3.1 2D Fourier version

We consider a simplified case where we have a 1D scanline of pixels and each pixel comes from a 1D integral, for example the integration of a linear light source or integration over time for motion blur. We call the scanline dimension x and the other one t. One version of the problem is

$$I(x) = \int f(x,t)w(t)\,\mathrm{d}t \quad (20)$$

where $w(t)$ is the weight over time (e.g. a hat corresponding to shutter time). A version with antialiasing is

$$I(x) = \int\int f \otimes \ell(x,t)w(t)\,\mathrm{d}t \quad (21)$$

where ℓ is a prefilter. For antialiasing, we need to first sample f and then perform a Monte-Carlo integration of the prefilter convolution (supersampling). The motion blur integral can also be seen as a convolution. For simplicity, we can note w'(x,t) the convolution kernel corresponding to both motion blur and antialiasing.

In practice, we first sample f through supersampling, and perform the convolutions on this sampled version.

$$I_S(x) = [(fS) \otimes w'](x) \quad (22)$$

Note that we started with a 2D integrand f and eventually slice it into a 1D image. In the Fourier domain, the multiplication becomes a convolution $\hat{S} \otimes \hat{f}$, the convolution a multiplication by $\hat{w}'$, and slicing is an integral over ω_t.

$$\hat{I_S}(\omega_x) = \int \left[(\hat{f} \otimes \hat{S})\hat{w'}\right](\omega_x, \omega_t)\,\mathrm{d}\omega_t \quad (23)$$

Sampling creates the usual replicas. The convolution by the prefilter/exposure kills high frequencies but the replicas inside that window remain.

What is the error for a given spatial frequency ω_x? It is the integral over ω_t of the aliases for that spatial frequency.

Note that negative interference can happen and reduce the effect of aliasing. This is similar to the 1D case where the DC term received the integral of the aliases.

Mitchell's spectrally optimal sampling This integration over ω_t made Mitchell [1991] advocate sampling patterns that avoid a cylinder of low spatial frequencies. However, note that only the aliases inside the (soft) window defined by $\hat{w}'$ contribute to the error. It is not different from traditional aliasing. This is why we respectfully disagree with his conclusions. The sampling patterm should just push frequency content outside the bandwidth of the antialiasing/shutter exposure filter.

The problem in Mitchell's argument is that he considers the spectrum of the integrand without the effect of sampling and looks at the effect of convolution and integration over that spectrum. However, sampling and these operations do not commute and one has to start with sampling to derive the correct spectra. This error is surprising because the previous section of his paper does contain the correct derivation of the sampled spectrum.

L2 error We can compute the L2 error over the image. According to Parseval, we can compute it directly in the Fourier domain:

$$\begin{aligned}
||I - I_S|| &= \int (I(x) - I_S(x))^2\,\mathrm{d}x && (24)\\
&= \int \left(\hat{I}(\omega_x) - \hat{I_S}(\omega_x)\right)^2 \mathrm{d}\omega_x && (25)\\
&= \int_{\omega_x}\left[\int_{\omega_t}\left((\hat{S'} \otimes f)\hat{w'}\right)(\omega_x,\omega_t)\,\mathrm{d}\omega_t\right]^2 \mathrm{d}\omega_x && (26)
\end{aligned}$$

The first integral is on values while the second one is on square. As a result, the total L2 error is smaller than by the variance of the aliases.

Interlude How does this relate to our simple 1D case where Monte-Carlo integration extracted the DC? Why do the higher ω_t matter now? This is because they are cross frequencies in both time and space, and they have an effect on the spatial frequencies of the final image.

Maybe we can look at it another way. At each x, the image value is the DC of S restricted to x multiplied by f, convolved by w'. This restriction of S to x corresponds to an integral of the spectrum $\hat{S}$. Furthermore, the DC gives us the value in the primal, but we are interested in the spectrum of the resulting image.

4 Discussion

Fourier analysis considers an infinite support, whereas Monte-Carlo integration is usually studied on a [01] interval. In particular, we observe that many quadrature schemes spend much resources (i.e. samples) around the extremities. This is for example the case with Gaussian quadrature, where the density of samples is higher around 0 and 1, and their weight is correspondingly smaller, roughly corresponding to importance sampling with higher importance at the boundaries.

In Fourier analysis, if we window the domain or periodize the function, we might also suffer from extraneous frequencies due to the lack of continuity at the boundary.

References

B , J. P. 2011. *Chebyshev and Fourier Spectral Methods*, 2nd ed., vol. 1. Dover Books on Mathematics, College Station, Texas.

C , R. L. 1986. Stochastic sampling in computer graphics. *ACM Transactions on Graphics 5*, 1 (Jan.), 51–72.

C -U , D., N , C. J. 2002. Sharp error bounds for the trapezoidal rule and simpson's rule. *Journal of Inequalities in Pure and Applied Mathematics 2*, 4.

D , F., H , N., S , C., C , E., S , F. X. 2005. A frequency analysis of light transport. *ACM Transactions on Graphics 24*, 3 (Aug.), 1115–1126.

E , K., T , Y.-T., H , N., D , F., R - , R. 2009. Frequency analysis and sheared reconstruction for rendering motion blur. *ACM Transactions on Graphics 28*, 3 (July), 93:1–93:13.

E , K., D , F., R , R. 2011. Practical filtering for efficient ray-traced directional occlusion. *SIGGRAPH Asia 2011* (Dec.).

E , K., H , F., D , F., R , R. 2011. Frequency analysis and sheared filtering for shadow light fields of complex occluders. *ACM Transactions on Graphics 30*, 2 (Apr.), 1–13.

G , W. 1997. *Numerical analysis: an introduction*. Birkhäuser Verlag, pub-BIRKHAUSER:adr.

M , D. P. 1991. Spectrally optimal sampling for distributed ray tracing. In *Computer Graphics (Proceedings of SIGGRAPH 91)*, vol. 25, 157–164.

O , M. J., F , E. 2001. On numerical solutions to one-dimensional integration problems with applications to linear light sources. *ACM Transactions on Graphics 20*, 4 (Oct.), 232–279.

R , R., H , P. 2001. A signal-processing framework for inverse rendering. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, 117–128.

S , C., S , K., D , F., H , N., S , F. X. 2009. Fourier depth of field. *ACM Trans. Graph. 28*, 2.

W , J. A. C. 2002. Numerical integration of periodic functions: A few examples. *The American Mathematical Monthly 109*, 1 (Jan.), 21–36.

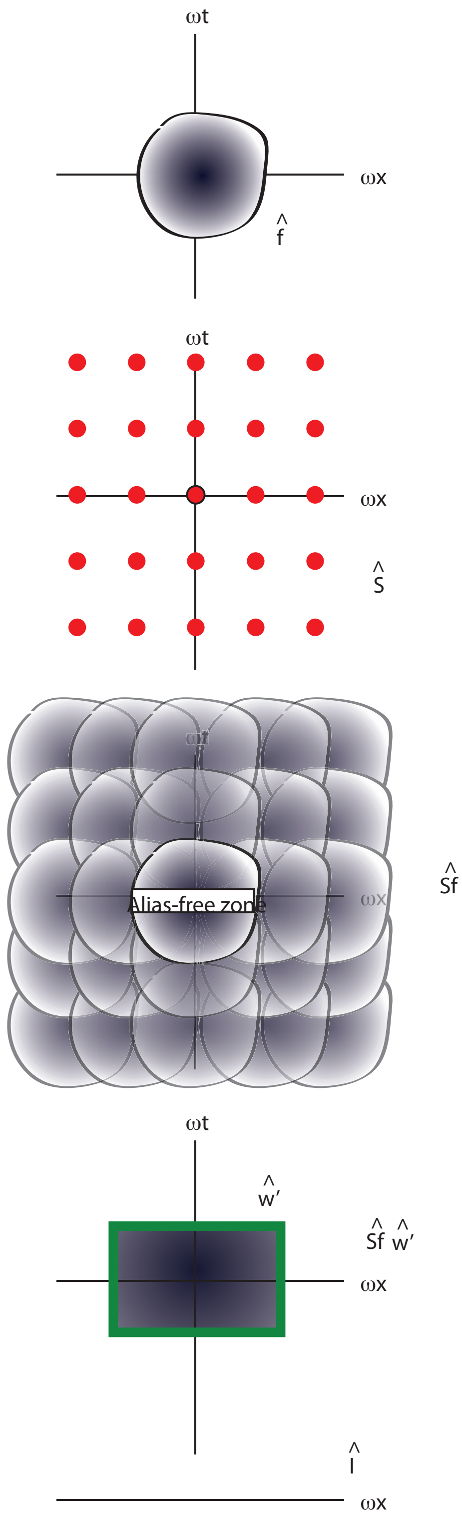

Figure 3: *Fourier-domain interpretation of supersampled image synthesis with motion blur.*