Basic Fourier integrals

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

August 3, 2014

1 Introduction

"The series $\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$ converges, and indeed uniformly, if $\sum(|a_n| + |b_n|)$ converges. Apart from this trivial case the convergence of trigonometric series is a delicate problem". A Zygmund, "Trigonometric Series", Volume 1, Cambridge University Press, 1959, page 4)

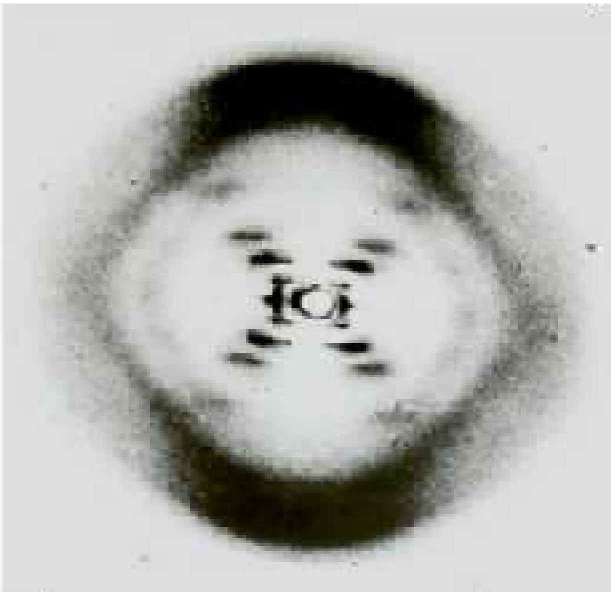
Fourier theory is a profound subject which is like a mathematical version of kikuyu grass with runners that go in every direction and eventually cover the entire ground. Learning it as a student only gives you the smallest of insights into the breadth of the subject but one has to start somewhere. In this paper what I have done is provide the essential building blocks of Fourier analysis in terms of the integral transform and in so doing I have followed the approach of Elias Stein and Rami Shakarchi in their superb Princeton Lecture Series [3]. There is a handful of basic properties which, when applied to the study of differential equations in particular, will take you a long way. To this end I have set out the basic properties with detailed proofs based upon the assumption that the functions inhabit Schwartz space. One can do Fourier theory without referring to Schwartz space and indeed this is done in many introductory courses, but ultimately you will need to visit Schwartz space. Indeed, as will be shown shortly, the fact that it is possible to get away without Schwartz space is instructive in itself because the theory still works. In engineering contexts this sort of detail is generally skated over but for those who want to really know why the theory works as well as it does you have to get your hands dirty with the nitty gritty of functional analysis. At one level Fourier theory is schizophrenic - it is possible to teach it with relatively little rigour even though analytical tripwires cover the entire ground, or one can adopt a highly analytical approach which can obscure the remarkable physical aspects of the theory. I have tried to steer a middle course and to do so I chosen a variant of the basic heat equation and the Black-Scholes partial differential equation as mechanisms to demonstrate how the theory works in detail. To give an idea of the sheer scope of Fourier theory here are some examples.

Typical applications of classical Fourier analysis are to:

• Frequency Modulation: Alternating current,radio transmission;
• Mathematics: Ordinary and partial differential equations, analysis of linear and nonlinear operators;Number Theory;
• Medicine: Electrocardiography, magnetic resonance imaging, biological neural systems;
• Optics and Fibre-Optic Communications: Lens design, crystallography, image processing;
• Radio,Television, Music Recording: Signal compression, signal reproduction, filtering;
• Spectral Analysis: Identification of compounds in geology, chemistry, biochemistry, mass spectroscopy;
• Telecommunications: Transmission and compression of signals, filtering of signals, frequency encoding.

For a contemporary application of Fourier theory and wavelet theory in the context of analysing climate change, see mathematical physicist John Baez's Azimuth Project: `http://johncarlosbaez.wordpress.com`. His website contains material which explains how Gabor transforms are used in the context of searching for patterns in the Southern Oscillation Index data. That index is relevant to El Nino and La Nina phenomena.

The discovery of the double helix structure of DNA is contained in the famous x-ray diffraction "photograph 51" generated by Rosalind Franklin and Maurice Wilkins and they used Fourier theory and Bessel functions to define the structure in a form that looked like this: $F_n = J_n(2\pi rR)\, e^{in(\phi+\frac{\pi}{2})}$ where $J_n(u)$ is the n^{th} order Bessel function of u [see [1]]. James Watson who, along with Francis Crick, shared the Nobel Prize for the discovery of DNA apparently sneaked a look at the X-ray photograph on Franlin's desk and realised that it implied a double helix structure.



Even more remarkable at one level is the use of Fourier theory in number theory. An important example is Timothy Gower's proof of Szemerédi's Theorem [19] which states the following:

Let k be a positive integer and let $\delta > 0$. There exists a positive integer $N = N(k, \delta)$ such that every subset of the set $1, 2, \ldots., N$ of size at least δN contains an arithmetic progression of length k.

The proof involves consideration of subsets of $\mathbb{Z}_N$ where N is prime and for a function $f : \mathbb{Z}_N \to \mathbb{C}$ and $r \in \mathbb{Z}_N$. We set:

$\hat{f}(r) = \sum_{s\in\mathbb{Z}_N} f(s)\,\omega^{-rs}$ where $\omega = e^{\frac{2\pi i}{N}}$. The function f is the discrete Fourier transform of f and is used widely in analytic number theory. Indeed in his paper [19] Gowers makes the point that four fundamental properties of the Fourier transform so defined are used repeatedly. Thus if the convolution is written in non-standard form as:

$f * g(s) = \sum_{t\in\mathbb{Z}_N} f(t)\,\overline{g(t-s)}$

then the following four fundamental identities hold:

$\widehat{(f * g)}(r) = \hat{f}(r)\,\overline{\hat{g}(r)}$

$\sum_r \hat{f}(r)\,\overline{\hat{g}(r)} = N \sum_s f(s)\,\overline{g(s)}$

$\sum_r |\hat{f}(r)|^2 = N\ \sum_s |f(s)|^2$

$f(s) = \frac{1}{N}\ \sum_r \hat{f}(r)\,\omega^{rs}$

The first identity tells us that convolutions transform to pointwise products, the second and third are Parseval's identities and the last is the inversion formula.

Beyond Fourier theory is wavelet analsyis and more. The applications that are being developed for wavelet analysis are very similar to those just listed. But the wavelet algorithms give rise to faster and more accurate image compression, faster and more accurate signal compression, and better denoising techniques that preserve the original signal more completely. The applications in mathematics lead, in many situations, to better and more rapid convergence results.

2 Background

At its most basic level we can guarantee the existence of the Fourier transform and its inverse with a simple test that doesn't require profound analysis. Thus if $\int_{-\infty}^{\infty} |f(x)|\,dx < \infty$ then $\mathcal{F}f$ and $\mathcal{F}^* f = \mathcal{F}^{-1} f$ exist and are continuous.

We can establish existence by noting that:

$|\mathcal{F}f(\xi)| = \left| \int_{-\infty}^{\infty} f(x)\,e^{-2\pi i x\xi}\,dx \right| \leq \int_{-\infty}^{\infty} |f(x)|\,|e^{-2\pi i x\xi}|\,dx = \int_{-\infty}^{\infty} |f(x)|\,dx < \infty$

Continuity follows from the following estimate for any ξ and ξ':

$|\mathcal{F}f(\xi) - \mathcal{F}f(\xi')| = \left| \int_{-\infty}^{\infty} f(x)\,e^{-2\pi i x\xi}\,dx - \int_{-\infty}^{\infty} f(x)\,e^{-2\pi i x\xi'}\,dx \right| = \left| \int_{-\infty}^{\infty} f(x)\,(e^{-2\pi i x\xi} - e^{-2\pi i x\xi'})\,dx \right|$
$\leq \int_{-\infty}^{\infty} |f(x)|\,|e^{-2\pi i x\xi} - e^{-2\pi i x\xi'}|\,dx$

Because $\int_{-\infty}^{\infty} |f(x)|\,dx < \infty$ we can take the limit $\xi' \to \xi$ so that $|e^{-2\pi i x\xi} - e^{-2\pi i x\xi'}| \to 0$ and we get $|\mathcal{F}f(\xi) - \mathcal{F}f(\xi')| \to 0$ as $\xi' \to \xi$. This shows that $\mathcal{F}f$ is continuous and we can replicate the argument to show that $\mathcal{F}^* f$ is continuous. The argument can be made rigorous with the usual ϵ, δ, N paraphenalia and there is plenty of that to come.

The fact that the Fourier transform is continuous is non-trivial. The classical electrical engineering function (which reflects a physical reality - just connect an oscilloscope to the appropriate circuit) of the "box" function or square wave $\Pi(\xi)$ which is 1 on the interval $[-\frac{1}{2}, \frac{1}{2}]$ and zero otherwise. It is discontinuous but its Fourier transform is: $\Pi(\xi) = \int_{-\infty}^{\infty} e^{-2\pi i x\xi}\, \Pi(x)\,dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-2\pi i x\xi}\,dx = \operatorname{sinc}\xi$, which is continuous. Note that $\int_{-\infty}^{\infty} |\,\Pi(x)\,|\,dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} 1\,dx = 1$ so that $\Pi \in \mathbb{L}^1(\mathbb{R})$. Now here's the problem. The sinc function does not satisfy the simple absolute integrability condition mentioned above. In fact, $\int_{-\infty}^{\infty} |\operatorname{sinc} x\,|\,dx = \infty$. If sinc x were absolutely integrable we could be assured that its Fourier transform - which is $\Pi(\xi)$ - exists and is continuous, but we know that Π is not continuous. One can establish that sinc x is not absolutely integrable by a careful estimation process which is instructive in that it shows that although $|\operatorname{sinc}\xi| = \left|\frac{\sin \pi\xi}{\pi\xi}\right| \to 0$ as $\xi \to \pm\infty$, the $\frac{1}{\xi}$ factor does not do so fast enough to get convergence of the integral.

Now it can be shown with some pretty subtle analysis that $\int_{-\infty}^{\infty} e^{-2\pi i x\xi} \operatorname{sinc} x\,dx = \begin{cases} 1 & \text{if } |\xi| < \frac{1}{2} \\ 0 & \text{if } |\xi| > \frac{1}{2} \end{cases}$

Note that you have two oscillatory processes going on in the integrand and they operate in a way to achieve cancellations and this is in itself a subtle process that is indicative of the field. It gets worse. Try to find the Fourier transform of something as basic as $\cos 2\pi x$ using the ideas developed above and you won't succeed notwithstanding the fact that a cosine signal or a sine signal underpins western civilisation! This problem is ultmately resolved with the theory of tempered distributions and you get $\frac{1}{2}\delta(\xi - 1) + \frac{1}{2}\delta(\xi + 1)$ where $\delta(x)$ is the Dirac delta "function". Indeed, if you pick up a book on spectroscopy, for instance, you will find something like the following (see [21]):

$\delta(t) = \int_{-\infty}^{\infty} e^{+2\pi i t s}\,ds = \mathcal{F}(\mathbf{1}) = \int_{-\infty}^{\infty} \cos(2\pi t s)\,ds$

Suffice it to say that the world of spectroscopy has not crumbled under the weight of such unrigorous formalism, yet the full justification for such matters does require quite a bit of work on generalised functions which is not the purpose of this paper. To read more on the details of generalised functions see [22].

In short the full theory of tempered distributions (see [20] or more specialised textbooks for more details) provides a rigorous foundation for all the "usual suspects" of the real world of electrical engineering and signal analysis and much more. What follows is a basic introduction to the characteristics of the Schwartz space to show the power of the concepts of "tempered distributions" and "generalised functions".

The fact that the $e^{-\pi x^2}$ is its own Fourier transform is one of the most fundamental and interesting facts of Fourier theory. The crux of the Heisenberg Uncertainty Principle is that if something is "localised" in one space it is "spread out" in another. In other words, in some general sense a function and its Fourier transform cannot be fundamentally localised. Moreover, the Gaussian $e^{-\pi x^2}$ (subject to some scaling) causes the product of momentum uncertainty and position uncertainty $\Delta p\, \Delta q$ to be minimised. [see [1] page 19]. The Fourier transform is a bijective mapping on the space of Schwartz functions (for a proof see [2] pages 140-142). Thus the Gaussian $e^{-\alpha\, x^2}$, where $\alpha > 0$, inhabits Schwartz space and is a fixed point of the space. If $\mathcal{F}$ represents the process of taking a Fourier transform of a function, $\mathcal{F}(e^{-\pi x^2}) = e^{-\pi \xi^2}$ where $\mathcal{F}(f)(x) = \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x \xi}\, dx$ while the inverse transform is given by $\mathcal{F}^*(\hat{f})(\xi) = \int_{-\infty}^{\infty} f(\xi)\, e^{2\pi i x \xi}\, d\xi$. Thus in the case of $e^{-\pi x^2}$ we have $\mathcal{F}^*(e^{-\pi \xi^2}) = e^{-\pi x^2}$ or $\mathcal{F}^* \circ \mathcal{F} = \mathcal{I}$. This is proved later. (Note that $\mathcal{F}^* = \mathcal{F}^{-1}$ to save a keystroke!).

The Schwartz space of functions comprises those functions that decay rapidly enough so that the basic manipulations of Fourier theory work "nicely". Indeed, once we are in Schwartz space we can do the mathematical equivalent of terrible things to small furry animals without getting arrested! More specifically, the Schwartz space on $\mathbb{R}$ (denoted by $\mathcal{S}(\mathbb{R})$) is the set of all indefinitely differentiable functions f so that f and all its derivatves $f^{(l)}$ are rapidly decreasing in the sense that:

$$\sup_{x \in \mathbb{R}} |x|^k\, |f^{(l)}(x)| < \infty \text{ for all } k, l \geq 0 \tag{1}$$

Some authors prefer to write (1) in the equivalent form:
$\sup_{x \in \mathbb{R}} (1 + |x|)^k\, |f^{(l)}(x)| < \infty$ for all $k, l \geq 0$

That this form is equivalent to (1) can be seen by noting that $(1 + |x|)^k$ is simply a polynomial each of whose terms satisfies (1) and so the finite sum of such terms will also satisfy (1).

This form immediately avoids any issues at $x = 0$ when performing estimates of the general form $\int_{-\infty}^{\infty} |f(x)|\, dx \leq \int_{-\infty}^{\infty} \frac{c_k\, dx}{(1+|x|)^k}$

Schwartz did not pull (1) out of the air but unfortunately many expositions of the concept do just that. Schwartz space $\mathcal{S}(\mathbb{R})$ sits between $C_0^{\infty}(\mathbb{R})$ and $\mathcal{L}^1(\mathbb{R})$ such that its functions are invariant under $\mathcal{F}$ ie you stay in Schwartz space under $\mathcal{F}$. If we suppose that $f(x)$ and $x\, f'(x)$ are both integrable then we can show that the Fourier transform $\hat{f}(\xi)$ is differentiable with respect to ξ and in fact:

$\frac{d\hat{f}}{d\xi} = \mathcal{F}(-2\pi i x f(x)\,)$

Using induction we can then establish that if $x^n \cdot f(x)$ is integrable for every integer $n > 0$, then the Fourier transform is an infinitely differentiable function. Going a bit further, if f is a continuously differentiable function such that both $f(x)$ and $f'(x)$ are integrable and such that $\lim_{|x| \to \infty} f(x) = 0$, then $\lim_{|\xi| \to \infty} \xi \cdot \hat{f}(\xi) = 0$

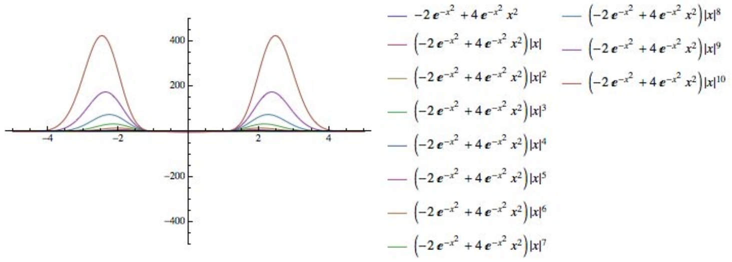
One can think of Schwartz functions either in terms of bounded products as in (1) or as limit property:

Limit property - boundedness equivalence

Thus if $f \in C^{\infty}(\mathbb{R})$ then $\lim_{|x|\to\infty}|x|^k\,|f^{(l)}(x)| = 0$ for all integers $k, l \geq 0$ if and only if $\sup_{x\in\mathbb{R}}\,|x|^k\,|f^{(l)}(x)| < \infty$ for all $k, l \geq 0$

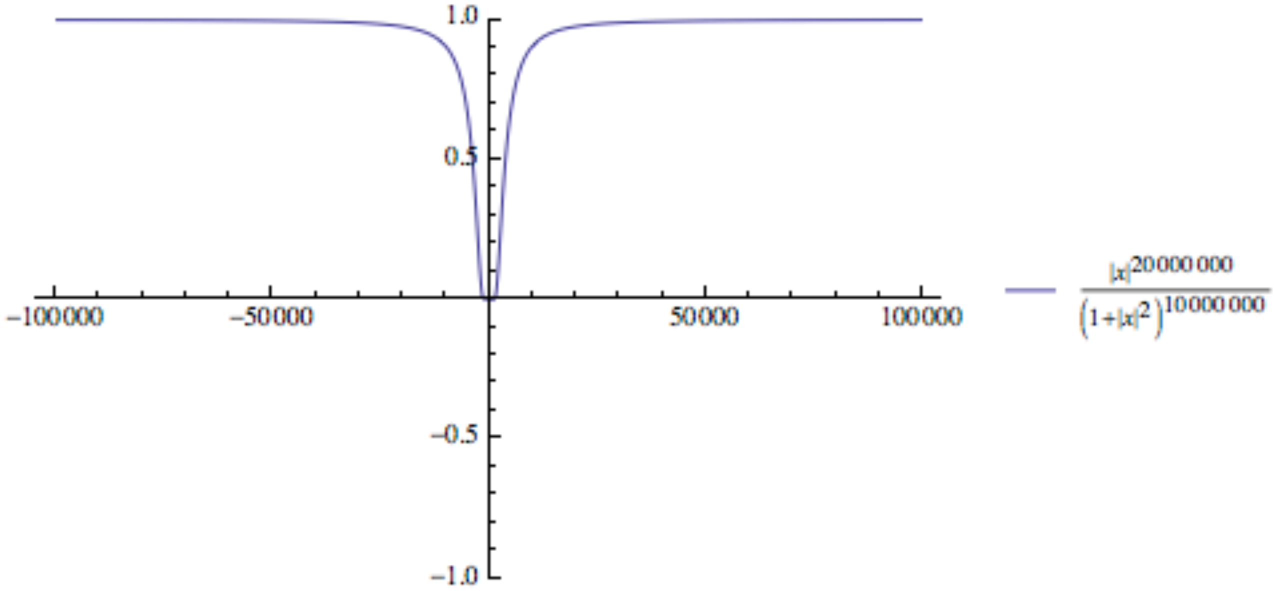
(1) contains a great deal of information since one can independently fix k and l. The Schwartz space therefore consists of smooth functions whose derivatives (including the function itself ie when $l = 0$) decay at infinity faster than any power. Not only do we have continuity of the Schwarz functions, even better, we have uniform continuity when we choose a closed interval. In (1) when $l = 0$ we have $\sup_{x\in\mathbb{R}}\,|x|^k\,|f(x)| < \infty$ for all k. This gives the hint of the basic form of $f(x)$ since we know that the exponential e^x grows faster than any power of x (and hence its inverse decays correspondingly faster) so it makes sense that a function such as e^{-x^2}, for instance, inhabits the Schwartz space. It is, of course, infinitely differentiable.

Clearly if $p(x)$ is any polynominal then $p(x)e^{-x^2}$ also lives in the Schwartz space because Schwartz functions fall off at infinity faster than the inverse of any polynomial. Looking at the derivatives of e^{-x^2} we see quickly that they are of the form $p(x)\,e^{-x^2}$, eg $f^{(2)}(x) = (4x^2 - 2)\,e^{-x^2}$. The following graph shows the boundedness and smoothness of $|x|^k\,|(4x^2 - 2)\,e^{-x^2}|$ for $0 \leq k \leq 10$.



It is worth noting here that $\mathcal{S}(\mathbb{R}) \subseteq \mathbb{L}^p(\mathbb{R}), \forall p\ 1 \leq p < \infty$. This follows since $|x|^2\,|f(x)| \to 0$ for $|x| \to \infty$. This implies that $|x|^{2p}\,|f(x)|^p$ is bounded, ie $\exists M > 0$ such that $|f(x)|^p \leq M|x|^{-2p} \leq \frac{M}{|x|^2}$ for $|x| \geq 1$. Thus $\int_{|x|\geq 1}|f(x)|^p\,dx \leq \int_{|x|\geq 1}\frac{M}{|x|^2}dx < \infty$

Note that $e^{-|x|}$ does not live in the Schwartz space simply because of the lack of differentiability at $x = 0$ notwithstanding the fact that it does fall off rapidly at infinity. Another example of a function that does not live in Schwartz space is $f(x) = \frac{1}{(1+|x|^2)^k}$ since $\frac{|x|^{2k}}{(1+|x|^2)^k}$ does not decay to zero for any k as $|x| \to \infty$. For instance, when $k = 10^6$ and $-100000 \leq x \leq 100000$ the product looks like this:



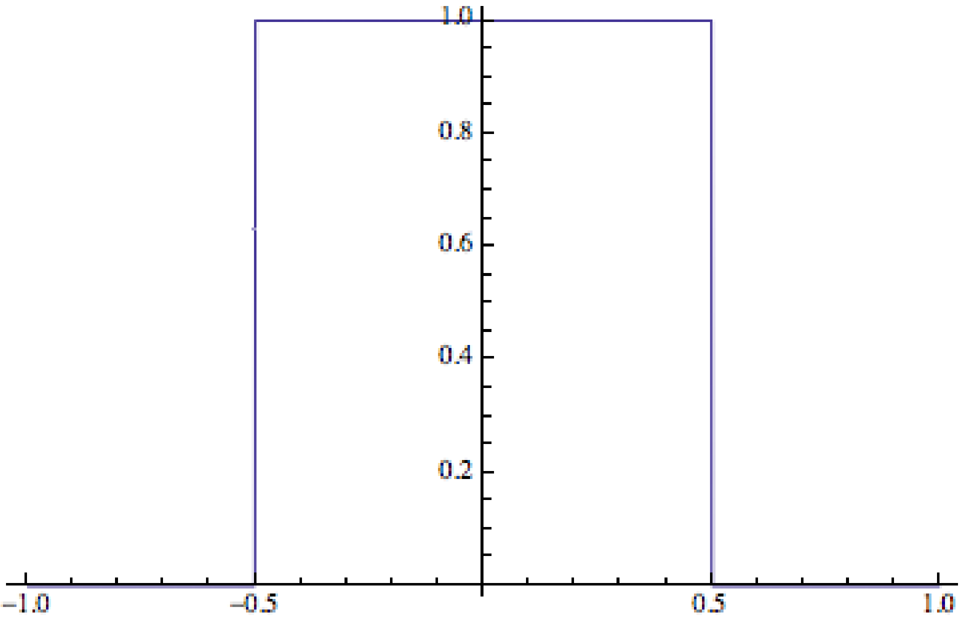
The function $f(x) = e^{-x^2} \sin(e^{x^2})$ is not a creature of Schwartz space because $f'(x)$ does not decay to zero as $|x| \to \infty$. This can be seen noting that $f'(x) = 2x\cos(e^{x^2}) - 2e^{-x^2}\, x \sin(e^{x^2})$ so that when x is large the derivative is dominated by the first term whose absolute value is unbounded. The second term goes to zero because $xe^{-x^2} \to 0$ as $x \to \infty$.

The Schwartz space concept emerged out of Laurent Schwartz's rigorous development of distribution theory during the Second World War (how he did this is a story in itself). In the 19^{th} century the proof techniques were centered around very intricate limit style proofs which inevitably had to show on a largely case by case that if you took some essentially arbitrary function $f(x)$ and you multiplied it by $e^{-2\pi i x\xi}$ and integrated, you actually got something that converged. Indeed, Fourier's original (and outrageous) assertion was that any periodic function could be represented by the series named after him. Thus Fourier believed that you could take his Fourier coefficients $c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx$ and you would get an infinite series $\sum_{n=-\infty}^{n=\infty} c_n\, e^{inx}$ which converges to $f(x)$ ie $f(x) = \sum_{n=-\infty}^{n=\infty} c_n\, e^{inx}$. This idea met resistance at the time (Lagrange and Poisson were strong opponents of his general approach) and led to many subtle and intricate limiting arguments as parts of the theory were verified during the 19^{th} century. Dirichlet was responsible for many important foundational results. Fourier was wrong in terms of the fine detail but his instincts were right - the Zygmund quotation given at the beginning of this paper captures the spirit of things - Fourier theory is a delicate matter. Even during the 20^{th} century there were doubts about aspects of the theory and it was not until 1966 that Lennart Carleson showed that for square integrable functions on $[0, 1]$ the Fourier partial sums converge pointwise almost everywhere. The proof is difficult to say the least. Carleson actually tried first to disprove the result and in an interview after he won the 2006 Abel Prize he elaborated on the history of the proof [16].

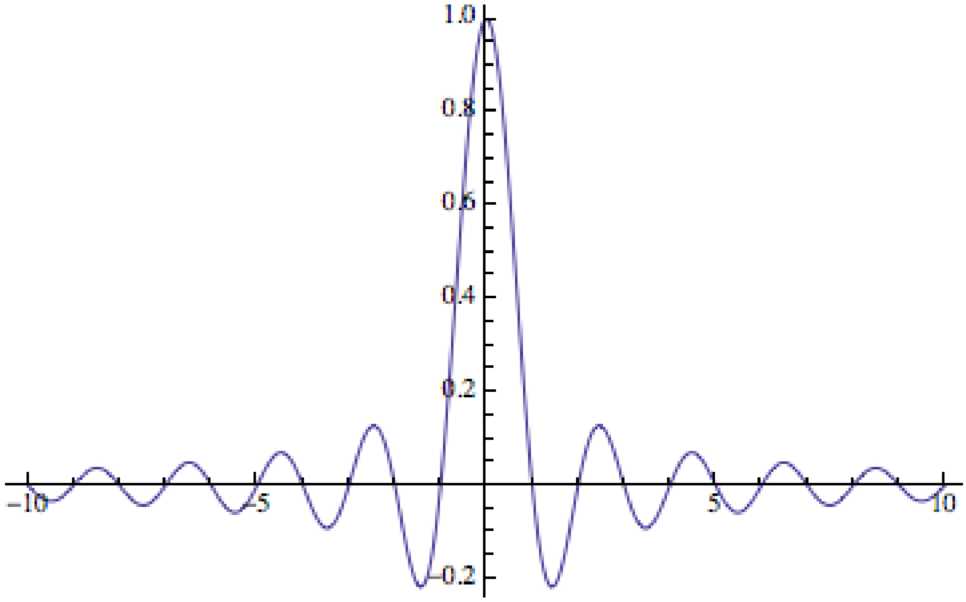
A classic example of how Fourier theory was approached in the 19^{th} century involves the treatment of the unit box function defined as follows:

$$f(x) = \begin{cases} 1 & \text{if } |x| < \frac{1}{2} \\ 0 & \text{if } |x| > \frac{1}{2} \end{cases} \tag{2}$$

Because f has a discontinuity at $x = \frac{1}{2}$ it was Dirichlet who came up with the idea of replacing the value of the function at the point of discontinuity by $\frac{f(x+0)+f(x-0)}{2}$. Thus in the case of the unit box function $f(\pm\frac{1}{2}) = \frac{1}{2}$. The Fourier transform of the unit box function is the sinc function:



transforms to:



Now f is clearly not in Schwartz space since it is discontinuous at $x = \pm\frac{1}{2}$ yet it has a bona fide Fourier transform. The conditions Dirichlet required for the basic theorem were that $\int_{-\infty}^{\infty}|f(x)|\,dx < \infty$ and that f and f' are piecewise continuous on every finite interval while at a point of discontinuity $f(x)$ is replaced by $\frac{f(x+0)+f(x-0)}{2}$. Under these assumptions, at each point where the function's one-sided derivatives exist, the function can be represented by:

$$f(x) = \frac{1}{\pi}\int_0^{\infty}\int_{-\infty}^{\infty} f(\xi)\,\cos[\alpha(\xi - x)]\,d\xi\,d\alpha \tag{3}$$

For the details of the intricate limiting arguments that Dirichlet used to establish (3) see Churchill's book (Chapter 6, [17]). I will not reproduce them here but they were the "bread and butter" of traditional Fourier theory courses.

Now if (3) is valid (and it is because f satisfies the hypotheses of the theorem) we should be able to show that $f(\pm\frac{1}{2}) = \frac{1}{2}$. Performing the integration we see that:

$$\begin{aligned}
f(x) &= \frac{1}{\pi}\int_0^{\infty}\int_{-\infty}^{\infty} f(\xi)\,\cos[\alpha(\xi - x)]\,d\xi\,d\alpha = \frac{1}{\pi}\int_0^{\infty}\int_{-\frac{1}{2}}^{\frac{1}{2}} \cos[\alpha(\xi - x)]\,d\xi\,d\alpha \\
&= \frac{1}{\pi}\int_0^{\infty}\Big[\frac{\sin[\alpha(\xi - x)]}{\alpha}\Big]_{-\frac{1}{2}}^{\frac{1}{2}}\,d\alpha = \frac{1}{\pi}\int_0^{\infty}\Big[\frac{\sin[\alpha(\frac{1}{2} - x)]}{\alpha} - \frac{\sin[\alpha(-\frac{1}{2} - x)]}{\alpha}\Big]\,d\alpha \\
&= \frac{1}{\pi}\int_0^{\infty} 2\,\frac{\sin\frac{\alpha}{2}\,\cos\alpha x}{\alpha}\,d\alpha
\end{aligned} \tag{4}$$

Hence from (4) we see that:

$$f(\frac{1}{2}) = \frac{1}{\pi}\int_0^\infty 2\,\frac{\sin\frac{\alpha}{2}\,\cos\frac{\alpha}{2}}{\alpha}\,d\alpha = \frac{1}{\pi}\int_0^\infty \frac{\sin\alpha}{\alpha}\,d\alpha = \frac{1}{\pi}\frac{\pi}{2} = \frac{1}{2} \tag{5}$$

Similarly, $f(-\frac{1}{2}) = \frac{1}{2}$. The proof that $\int_0^\infty \frac{\sin\alpha}{\alpha}\,d\alpha = \frac{\pi}{2}$ can be found in [17, pages 85-86] which I have set out in expanded form in the Appendix.

In broad terms with the Schwartz space of functions, the focus is on establishing the decay characteristics of the functions so that they decay sufficiently fast to damp out the intrinsic oscillatory behaviour of the exponential term in the integral. Once you define the appropriate rate of decay one can then perform generic limiting arguments with a degree of relative simplicity. It quickly becomes possible to say with confidence things like "This will be small because f is rapidly decreasing" without resort to tedious ϵ, N style arguments. Indeed, you can compress many steps in an otherwise detailed proof with such broad references and the chances are you will be ok! Stein and Shakarchi ([2] pages 131-134) briefly set out the foundation for the Schwartz space approach by first considering the space of "moderately" decreasing functions. These functions are assumed to be continuous on $\mathbb{R}$ and there exists a constant $A > 0$ such that $|f(x)| \leq \frac{A}{1+x^2}$ for all $x \in \mathbb{R}$.

The problem of convergence has practical dimensions in the context of quantum physics because Feynman's path integral which arose from Feynman's investigation of an integral that looked like this: $\int_{\text{all space}} e^{\frac{i\epsilon}{\hbar}\frac{m(x-y)^2}{2\epsilon^2}}\,\Psi(y,t)\,\frac{dy}{A}$ where $\Psi(y,t)$ is the wave function, reflected the basic characteristics of Fourier-style integrals.

In this paper I prove some basic facts which are essential building blocks for more complex problems. For instance, in the theory of the heat equation you have to evaluate a somewhat daunting looking integral of the form $\int_{-\infty}^{\infty} e^{(-4\pi^2\,\xi^2+(1-a)2\pi i\xi)t}\,e^{2\pi i\xi v}\,d\xi$ and some of the techniques explored below are relevant to solving that type of problem (see below for how this particular integral is solved). Later in this paper I deal with the heat equation and its relevance to the solution to the Black-Scholes equation.

3 Physical considerations

The Gaussian kernel (and the Dirac function for that matter) are not mere mathematical abstractions invented for the delectation of analysts. In fact physics drove the development of the Dirac function in particular. In advanced physics textbooks there are derivations of the Maxwell equations using *microscopic* rather than macroscopic principles eg see [section 6.6 of [14]]. If you follow the discussion in that book by Jackson you will see that for dimensions large compared to 10^{-14}m the nuclei can be treated as point systems which give rise to the microscopic Maxwell equations:

$$\nabla \cdot \mathbf{b} = 0, \quad \nabla \times \mathbf{e} + \frac{\partial \mathbf{b}}{\partial t} = 0, \quad \nabla \cdot \mathbf{e} = \frac{\eta}{\epsilon_0}, \quad \nabla \times \mathbf{b} - \frac{1}{c^2}\frac{\partial \mathbf{e}}{\partial t} = \mu_0 \mathbf{j}$$

Here **e** and **b** are he microscopic electric and magnetic fields and η and **j** are the microscopic charge and current densities. A question arises as to what type of averaging of the microscopic fluctations is appropriate and Jackson says that "at first glance one might think that averages over both space and time are necessary. But this is not true. Only a spatial averaging is necessary" [[14] , page 249] Briefly the broad reason is that in any region of macroscopic interest there are just so many nuclei and electrons so that the spatial averaging "washes" away the time fluctuations of the microscopic fields which are essentially uncorrelated at the relevant distance $(10^{-8}m)$.

The spatial average of $F(\mathbf{x}, t)$ with respect to some test function $f(\mathbf{x})$ is defined as:

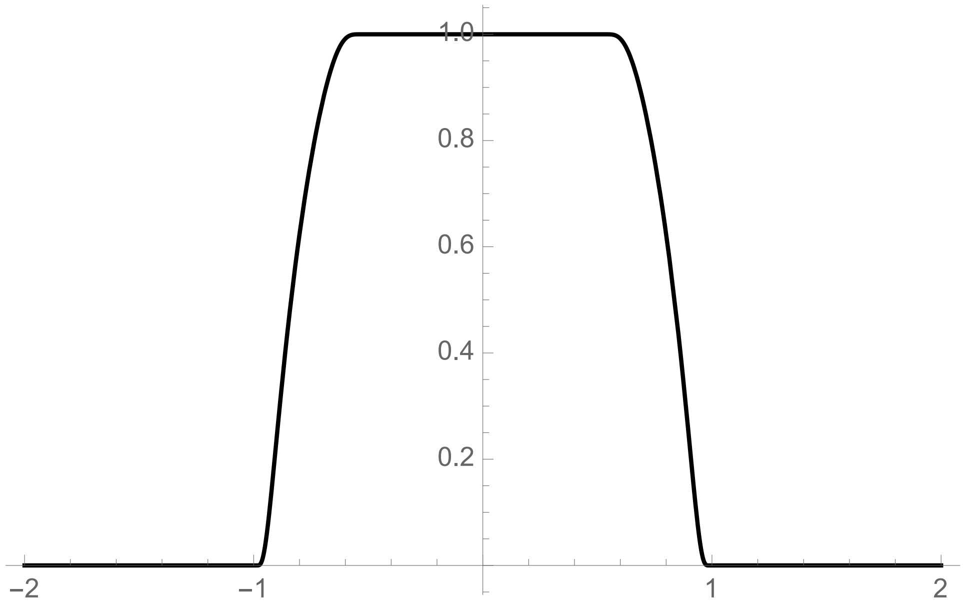
$\langle F(\mathbf{x}, t)\rangle = \int F(\mathbf{x} - \mathbf{x}', t)\, f(\mathbf{x}')\, d^3x'$ where $f(\mathbf{x})$ is real and non-zero in some neighbourhood of $\mathbf{x} = \mathbf{0}$ and is normalised to 1 over all space. It is reasonable to expect that $f(\mathbf{x})$ is isotropic in space so that there are no directional biases in the spatial averages. Jackson gives two examples as follows:

$$f(\mathbf{x}) = \begin{cases} \frac{3}{4\pi R^3}, & r < R \\ 0, & r > R \end{cases}$$

and

$$f(\mathbf{x}) = (\pi R^2)^{-\frac{3}{2}} e^{-\frac{r^2}{R^2}}$$

The first example is an average of a spherical volume with radius R but it has a discontinuity at $r = R$. Jackson notes that this "leads to a fine-scale jitter on the averaged quantities as a single molecule or group of molecules moves in or out of the average volume" [[14], page 250]. This particular problem is eliminated by a Gaussian test function "provided its scale is large compared to atomic dimensions" [[14] , p.250]. Luckily all that is needed is that the test function meets general continuity and smoothness properties that yield a rapidly converging Taylor series for $f(\mathbf{x})$ at the level of atomic dimensions. Thus the Gaussian plays a fundamental role in the calculations presented by Jackson concerning this issue. The article upon which Jackson's comments are based is that of G Russkaoff in [15]. Jackson gives as an example of the type of Gaussian test function something like the following [see [14], p.250 Figure 6.1]:



The *Mathematica* code to generate this graph is as follows:

```
g[x_ /; 0 < x < 1/2] := e^((-1/(2 x)) e^(1/(2 x-1)));
g[x_ /; x ≤ 0] := 0;
g[x_ /; x ≥ 1/2] := 1;

c = Plot[g[1 + x] * g[1 - x], {x, -2, 2}, PlotStyle → Black]
```

Note that the function in the above graph is infinitely differentiable and is bounded.

4 Some building blocks

Proving the various fundamental building blocks of Fourier transforms on Schwartz functions necessarily involves proving that an infinite integral converges. The analysis thus revolves around the "hump" and "tails" of the integral estimates. Broadly, the tails will be small because the Schwartz function $|f(x)|$ will be rapidly decreasing for large values of x. To get the hump sufficiently small one will usually need to use the fact that the Schwartz functions are uniformly continuous on closed, bounded intervals. The details will become clearer in the examples which follow.

In what follows we assume that f inhabits the Schwartz space. The Fourier transform of f is:

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x\xi}\, dx \tag{6}$$

There are other ways to define the Fourier transform. Other possibilities are $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-i\xi x}\, dx$, $\int_{-\infty}^{\infty} f(x)e^{-i\xi x}\, dx$ and $\int_{-\infty}^{\infty} f(x)e^{+i\xi x}\, dx$. In quantum physics if $\psi(x)$ is a one dimensional wave function, its Fourier transform $\overline{\psi}(p)$ is conventionally defined as $\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x)\, e^{\frac{-ipx}{\hbar}}\, dx$ where $\hbar = \frac{h}{2\pi}$ and h is Planck's constant. The inverse transform is then $\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \bar{\psi}(p)\, e^{\frac{ipx}{\hbar}}\, dp$ (see [3], p.1462]).

The "magic" of Fourier theory enables us to get back to $f(x)$ as follows:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{2\pi i x\xi}\, d\xi \tag{7}$$

That one can do this is due to the Fourier inversion theorem which is proved later.

Let's start with the Fourier transform of $f^{'}(x)$ because without this little "engine" we couldn't do anything very useful with differential equations. It turns out to be $2\pi i\xi\, \hat{f}(\xi)$. Thus under Fourier transforms differentiation transforms simply as a product of a multiple of the transformed variable and the function's Fourier transform. So if you have a partial differential equation such as:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial t^2} \tag{8}$$

where $u(x, t)$ is a function of spatial and temporal dimensions ((8) is in fact the basic heat equation) there are advantages in taking the Fourier transform of both sides with respect to the spatial dimension. Using the rule (twice in the RHS of (8)) we have not yet proved you get:

$$\frac{\partial \hat{u}(\xi, t)}{\partial t} = -4\pi^2 \xi^2\, \hat{u}(\xi, t) \tag{9}$$

Note that the Fourier transform of $\frac{\partial u(x,t)}{\partial t}$ with respect to x is:

$$\int_{-\infty}^{\infty} \frac{\partial(x, t)}{\partial t}\, e^{-2\pi i x\xi}\, dx = \frac{\partial}{\partial t} \int_{-\infty}^{\infty} u(x, t)\, e^{-2\pi i x\xi}\, dx = \frac{\partial}{\partial t}\hat{u}(\xi, t) \tag{10}$$

That this is the case follows from the rules in relation to differentiating under the integral sign (Leibnitz's rule) - remember we are in Schwartz space so we can do terrible things to small furry animals with impunity! More details can be found in the Appendix.

Since (9) is just a garden variety differential equation in t if we fix ξ, you will get $\hat{u}(\xi, t) = A(\xi)\, e^{-4\pi^2\xi^2 t}$. To see this use the integrating factor $e^{4\pi^2\xi^2 t}$ as follows:

$$\frac{\partial}{\partial t}\{\hat{u}(\xi, t)\, e^{4\pi^2\xi^2 t}\} = e^{4\pi^2\xi^2 t}\Big\{\frac{\partial \hat{u}(\xi, t)}{\partial t} + 4\pi^2\xi^2\, \hat{u}(\xi, t)\Big\} = 0 \tag{11}$$

Hence $\hat{u}(\xi, t)\, e^{4\pi^2\xi^2 t} = A(\xi)$ (ie some function independent of t) and so $\hat{u}(\xi, t) = A(\xi)\, e^{-4\pi^2\xi^2 t}$. Because (8) will involve initial conditions we need to also take the Fourier transform of those conditions and we ultimately end with up with:

$$\hat{u}(\xi, t) = g(\xi)\, e^{-4\pi^2\xi^2 t} \tag{12}$$

This may not look like it helps but it does because on the RHS of (12) we have a Gaussian and we know that the Fourier transform of a Gaussian is another Gaussian (generally with a scale factor) so the RHS of (12) is essentially the product of two Fourier transforms and it is a basic result of Fourier theory that the Fourier transform of a convolution of f and g is $\hat{f}(\xi)\, \hat{g}(\xi)$ ie:,

$$\hat{u}(\xi, t) = (\widehat{f * g})(\xi) = \int_{-\infty}^{\infty} f(\xi - y)g(y)\, e^{-2\pi i y\xi}\, dy = \hat{f}(\xi)\, \hat{g}(\xi) \tag{13}$$

To get back to $u(x, t)$ we use Fourier inversion: $\mathcal{F}^*\Big\{(\widehat{f * g})(\xi)\Big\} = \mathcal{F}^*\Big\{\mathcal{F}\Big(u(x, t)\Big)\Big\} = u(x, t)$. Note that $\hat{u}(\xi, t) = \mathcal{F}\Big\{u(x, t)\Big\}$. In the case of the Black-Scholes equation which has its roots in the heat equation, using these Fourier transform techniques on the partial differential equation which looks like this for $0 < t < T$:

$$\frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} - rV = 0 \tag{14}$$

you get a solution that looks like this:

$$V(s, t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^{\infty} e^{\frac{-(\log(\frac{s}{s^*}) + (r - \frac{\sigma^2}{2})(T-t))^2}{2\sigma^2(T-t)}}\, F(s^*)\, \frac{ds^*}{s^*} \tag{15}$$

5 Proofs of some basic transform properties

1. $\mathbf{f}^{'}(\mathbf{x}) \rightarrow \mathbf{2\pi i\xi\,\hat{f}}(\xi)$

To prove that $f^{'}(x) \rightarrow 2\pi i\xi\,\hat{f}(\xi)$ just apply the definition of the Fourier transform and use integration by parts. Thus we have:

$$\int_{-n}^{n} f^{'}(x)\,e^{-2\pi ix\xi}\,dx = \left[f(x)e^{-2\pi ix\xi}\right]_{-n}^{n} + 2\pi i\xi \int_{-n}^{n} f(x)\,e^{-2\pi ix\xi}\,dx \tag{16}$$

Now as $n \rightarrow \infty$, the first term in (16) goes to zero. To see this we have to go back to the properties of Schwartz functions.

From the definition of Schwartz space we can take $l = 0$ so that $\sup_{x\in\mathbb{R}} |x|^k\,|f(x)| < \infty$ for all $k \geq 0$. Because this is a global bound, this means that for any $n \neq 0$ we care to choose, $\exists s > 0$ such that $|n|^k\,|f(\pm\,n)| < s$ for all $k \geq 0$. Thus if we are given any $\epsilon > 0$ we can choose k large enough so that $\frac{s}{|n|^k} < \epsilon$, a property which holds for any n we choose. So $|f(\pm\,n)| < \frac{s}{|n|^k} < \epsilon$ for this k. In other words $|f(\pm\,n)| \rightarrow 0$ as $n \rightarrow \infty$. Note that at $n = 0$ we know that f is continuous and hence is bounded on any closed interval so that it does not blow up.

Applying this estimate we see that:

$$\left|f(n)\,e^{-2\pi\,i\,\xi\,n} - f(-n)\,e^{2\pi\,i\,\xi\,n}\right| \leq |f(n)| + |f(-n)| < \epsilon + \epsilon \tag{17}$$

This establishes that $\left[f(x)\,e^{-2\pi\,i\,\xi\,x}\right]_{-n}^{n} \rightarrow 0$ as $n \rightarrow \infty$ and so we get the result we wanted. Alternatively, if one uses the limit form of the Schwartz space definition this last estimate comes immediately since $\lim_{|x|\rightarrow\infty}|f(x)| \rightarrow 0$ (setting $k = 0$ in the definition).

Note that the definition of Schwartz space functions allows us to inductively assert that $a_1|x|\,|f(x)| < s_1$, $a_2|x|^2\,|f(x)| < s_2$,...,$a_n|x|^n\,|f(x)| < s_n$ so that $|P(x)|\,|f(x)| < S$ where $P(x)$ is a polynomial. This means that f decays faster at infinity than the inverse of any polynomial. This fact is used in some later estimates.

2. $\mathbf{-2\pi ix\,f(x)} \rightarrow \frac{\mathbf{d}}{\mathbf{d}\xi}\,\mathbf{\hat{f}}(\xi)$

To prove this we need to establish that $\hat{f}$ is actually differentiable (which ought to be the case given that f lives in the Schwartz space) and actually find the derivative. We start with this difference which is the definition of the derivative:

$$\frac{\hat{f}(\xi+h)-\hat{f}(\xi)}{h} - (\widehat{-2\pi ix}\,f)(\xi) = \int_{-\infty}^{\infty}\Big[\frac{f(x)\,e^{-2\pi ix(\xi+h)} - f(x)\,e^{-2\pi ix\xi}}{h} + 2\pi ix\,f(x)e^{-2\pi ix\xi}\Big]dx$$

$$= \int_{-\infty}^{\infty} f(x)\,e^{-2\pi ix\xi}\Big[\frac{e^{-2\pi ixh}-1}{h} + 2\pi ix\Big]\,dx = I \quad (18)$$

It is common in dealing with estimation problems such as that posed by (18) to break the integral up into three parts as follows:$(-\infty,-N)$, $[-N,N]$ and (N,∞) and to use relevant properties such as rapid decrease in the case of Schwartz space functions perhaps in combination with continuity, uniform continuity or differentiabilty as appropriate to ensure that the estimates are sufficiently small. In the case of (18) we can use the rapid decrease of $f(x)$ and $xf(x)$ to make the tails of the integral in (18) sufficiently small. The tails are usually the easiest part of the estimation problem whereas the middle "rump" tends to require more subtle estimates. Note here that if f is in $\mathcal{S}(\mathbb{R})$ so is $xf(x)$ which follows directly from the definition in (1) where one could simply write $\sup_{x\in\mathbb{R}} |x|^{k-1}\ |x\,f^{(l)}(x)| < \infty$ for all $k \geq 1, l \geq 0$.

An important property of f is that $\int_{|x|>N} |x|\,|f(x)|\,dx \to 0$ as $N \to \infty$. To see this we note that because f is in Schwartz space it is bounded and rapidly decreasing so $\exists B > 0$ such that $|x|^3\,|f(x)| < C$ for some constant C, for all $|x| > 1$. Thus, $|x|\,|f(x)| < \frac{C}{|x|^2}$ and so for any $\epsilon > 0$ we can find an N such that $\int_{|x|>N} |x|\,|f(x)|\,dx < \int_{|x|>N} \frac{C}{|x|^2}\,dx = \frac{2C}{N} < \epsilon$.

Our two (the tails are compressed into one) integrals are:

$$I_1 = \int_{|x|>N} f(x)\,e^{-2\pi ix\xi}\Big[\frac{e^{-2\pi ixh}-1}{h} + 2\pi ix\Big]\,dx \quad (19)$$

$$I_2 = \int_{-N}^{N} f(x)\,e^{-2\pi ix\xi}\Big[\frac{e^{-2\pi ixh}-1}{h} + 2\pi ix\Big]\,dx \quad (20)$$

Thus:

$$I = I_1 + I_2 \quad (21)$$

so that:

$$|I| \leq |I_1| + |I_2| \quad (22)$$

Now:

$$|I_1| \leq \int_{|x|>N} \Big| f(x)\, e^{-2\pi i x\xi} \Big[\frac{e^{-2\pi i x h} - 1}{h} + 2\pi i x \Big] \Big| \, dx \leq \int_{|x|>N} |f(x)| \Big| \Big[\frac{e^{-2\pi i x h} - 1 + 2\pi i x h}{h} \Big] \Big| \, dx$$

$$\leq 2\pi \int_{|x|>N} |x|\, |f(x)| \Big\{ \frac{|e^{-2\pi i x h} - 1|}{|2\pi i x h|} + 1 \Big\} \, dx \quad (23)$$

In relation to (23) we know that because of the rapid decrease of f we can make $\int_{|x|\geq N} |f(x)|\, |x| \, dx < \epsilon$ because $x f(x)$ is rapdily decreasing. To make $\int_{|x|>N} |x|\, |f(x)| \Big\{ \frac{|e^{-2\pi i x h}-1|}{|2\pi i x h|} + 1 \Big\} \, dx$ small enough we note that for any real θ the following inequality holds: $|e^{i\theta} - 1| \leq |\theta|$. That this is the case can be seen as follows:

$$|e^{i\theta} - 1| = |e^{\frac{i\theta}{2}} e^{\frac{i\theta}{2}} - e^{\frac{i\theta}{2}} e^{-\frac{i\theta}{2}}| = |e^{\frac{i\theta}{2}}|\, |e^{\frac{i\theta}{2}} - e^{-\frac{i\theta}{2}}| = |2 \sin \frac{\theta}{2}| \leq 2 \frac{|\theta|}{2} = |\theta| \quad (24)$$

Thus, $\frac{|e^{-2\pi i x h}-1|}{|2\pi i x h|} + 1 \leq 1 + 1 = 2$

Note here that this is a more refined bound on $|e^{i\theta} - 1|$ than the obvious one of 2 which would lead to an argument about the relative decay of $|x|\, |f(x)|\, (\frac{2}{2\pi\, |x|\, |h|} + 1)$.

Thus we see that:

$|I_1| = 2\pi \int_{|x|>N} |x|\, |f(x)| \, \Big\{ \frac{|e^{-2\pi i x h}-1|}{|2\pi i x h|} + 1 \Big\} \, dx \leq 4\pi \int_{|x|>N} |x|\, |f(x)| \, dx < 4\pi\epsilon$ using the fact that $\int_{|x|\geq N} |f(x)|\, |x| \, dx < \epsilon$.

We now come to $|I_2|$ which requires a slightly different line of attack because the interval contains the origin and even though f will be bounded at the origin we need to be sure that the integral is still sufficiently small when N is large. If we fix x and let $g(h) = e^{-2\pi i x h}$ we know that $g^{'}(h) = -2\pi i x\, e^{-2\pi i x h}$ and so $g^{'}(0) = -2\pi i x$. Note that $g(0) = 1$. In fact there is some $h_0 > 0$ such that for $|h| < h_0$ we have:

$$\Big| \frac{e^{-2\pi i x h} - 1}{h} - (-2\pi i x) \Big| = \Big| \frac{e^{-2\pi i x h} - 1}{h} + 2\pi i x \Big| < \frac{\epsilon}{2N} \quad (25)$$

The reason that (25) holds is that $\frac{e^{-2\pi i x h}-1}{h} + 2\pi i x$ converges uniformly to zero as $h \to 0$ for all $x \in [-N, N]$. In essence it is $\frac{g(h)-g(0)}{h} - g^{'}(0)$ which ought to converge uniformly to zero as $h \to 0$ for any $x \in [-N, N]$ because the derivative at issue is continuous on the compact interval $[-N, N]$ and hence uniformly continuous. This can be shown in more detail as follows. Fix $x \in [-N, N]$ and take $|h| < |h_1| < |h_0|$, and note that by the mean value theorem we can find $|h^{'}| < |h|$ such that $\frac{|e^{-2\pi i x h}-1|}{|2\pi x h|} = |2\pi x h^{'}|$ and similarly there is a $|h^{''}| < |h_1|$ such that such that $\frac{|e^{-2\pi i x h_1}-1|}{|2\pi x h_1|} = |2\pi x h^{''}|$. Then

$$\Big|\frac{e^{-2\pi ixh}-1}{h}+2\pi ix-\Big\{\frac{e^{-2\pi ixh'}-1}{h'}+2\pi ix\Big\}\Big| = \Big|\frac{e^{-2\pi ixh}-1}{h}-\Big\{\frac{e^{-2\pi ixh'}-1}{h'}\Big\}\Big|$$
$$\leq \Big|\frac{e^{-2\pi ixh}-1}{h}\Big| + \Big|\frac{e^{-2\pi ixh'}-1}{h'}\Big| = |2\pi x|\,\frac{|e^{-2\pi ixh}-1|}{|2\pi xh|} + |2\pi x|\,\frac{|e^{-2\pi ixh'}-1|}{|2\pi xh'|}$$
$$\leq |2\pi x|\,|2\pi xh^{'}| + |2\pi x|\,|2\pi xh^{''}| \leq 4\pi^2x^2(|h^{'}|+|h''|) < 4\pi^2N^2\,2|h_0| = 8\pi^2N^2\,|h_0| < \epsilon \quad (26)$$

if $|h_0| < \frac{\epsilon}{8\pi^2N^2}$. This establishes the uniform continuity. It is worth noting the subtlety of the estimate since we know from (24) that $\frac{|e^{i\theta}-1|}{|\theta|} \leq 1$ which is too big a bound to get the uniform continuity estimate to work. It is also worth noting that $\frac{|e^{i\theta}-1|}{|\theta|} = |\text{sinc}\frac{\theta}{2}|$ which is uniformly continuous: `http://www.gotohaggstrom.com/Uniform%20continuity%20of%20sinc%20x.pdf`

Thus we have:

$$|I_2| = \Big|\int_{-N}^{N} f(x)\,e^{-2\pi ix\xi}\Big[\frac{e^{-2\pi ixh}-1}{h}+2\pi ix\Big]dx\Big| \leq \int_{-N}^{N}\Big|f(x)\Big|\Big|\Big[\frac{e^{-2\pi ixh}-1}{h}+2\pi ix\Big]\Big|dx$$
$$\leq \int_{-N}^{N} B\frac{\epsilon}{2N}dx = B\epsilon \quad (27)$$

since $\exists B > 0$ such that $|f(x)| < B$ on $[-N, N]$. Putting the two estimates together we have:

$$|I| \leq\ 4\pi\epsilon + B\epsilon < C\,\epsilon \quad (28)$$

Hence the result follows. Note that Property 2 is a form of differentiation under the integral sign since it says $\frac{d}{d\xi}\int_{-\infty}^{\infty} f(x)\,e^{-2\pi ix\xi}\,dx = \int_{-\infty}^{\infty}(-2\pi ix)\,f(x)\,e^{-2\pi ix\xi}\,dx$.

3. $\mathbf{f(x+h)} \rightarrow \mathbf{\hat{f}}(\xi)\,\mathbf{e^{2\pi ih\xi}}$ for $h \in \mathbb{R}$

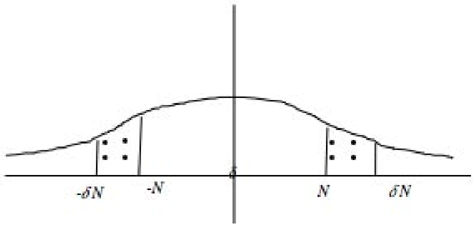
To prove this we note that $\mathcal{F}(f(x+h) = \int_{-\infty}^{\infty} f(x+h)e^{-2\pi ix\xi}\,dx$ and then make the substitution $x' = x+h$ so that $\int_{-\infty}^{\infty} f(x+h)e^{-2\pi ix\xi}\,dx = \int_{-\infty}^{\infty} f(x')e^{-2\pi i(x'-h)\xi}\,dx' = \hat{f}(\xi)\,e^{2\pi ih\xi}$.

4. $\mathbf{f(x)\,e^{-2\pi ixh}} \rightarrow \mathbf{\hat{f}(\xi+h)}$ for $h \in \mathbb{R}$

$\mathcal{F}(f(x)e^{-2\pi xh}) = \int_{-\infty}^{\infty} f(x)\,e^{-2\pi ixh}\,e^{-2\pi ix\xi}\,dx = \int_{-\infty}^{\infty} f(x)\,e^{-2\pi i(x+h)\xi}\,dx = \hat{f}(\xi+h)$

5. $\mathbf{f(\delta x)} \rightarrow \delta^{-1}\,\mathbf{\hat{f}}(\delta^{-1}\xi)$ for $\delta > 0$

Consider $\Big|\int_{-N}^{N} f(\delta x)\,e^{-2\pi ix\xi}\,dx - \delta^{-1}\int_{-N}^{N} f(x)\,e^{-2\pi ix\frac{\xi}{\delta}}\,dx\Big|$. Without loss of generality we can assume $\delta > 1$. If $0 < \delta \leq 1$ the limits of the relevant intervals are reversed.



$$
\begin{aligned}
\Big| \int_{-N}^{N} f(\delta x)\, e^{-2\pi i x\xi}\, dx - \frac{1}{\delta} \int_{-N}^{N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx \Big| \\
&= \Big| \frac{1}{\delta} \int_{-\delta N}^{\delta N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx - \frac{1}{\delta} \int_{-N}^{N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx \Big| \\
&= \frac{1}{\delta} \Big| \int_{-\delta N}^{\delta N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx - \int_{-N}^{N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx \Big| \\
= \frac{1}{\delta} \Big| \int_{-\delta N}^{-N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx + \int_{-N}^{N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx + \int_{N}^{\delta N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx - \int_{-N}^{N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx \Big| \\
&= \frac{1}{\delta} \Big| \int_{-\delta N}^{-N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx + \int_{N}^{\delta N} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx \Big| \\
&\leq \frac{1}{\delta} \left\{ \int_{-\delta N}^{-N} |f(x)|\, dx + \int_{N}^{\delta N} |f(x)|\, dx \right\} \leq \frac{1}{\delta} \left\{ \int_{-\delta N}^{-N} \frac{s}{|x|^k}\, dx + \int_{N}^{\delta N} \frac{s}{|x|^k}\, dx \right\} \\
&\leq \frac{2sN(\delta - 1)}{\delta N^k} = \frac{2s(\delta - 1)}{\delta N^{k-1}} \to 0 \text{ as } N \to \infty \qquad (29)
\end{aligned}
$$

Note here that we have used the same logic as applied in the lead up to (17).

What (29) shows is that $\int_{-\infty}^{\infty} f(\delta x)\, e^{-2\pi i x\xi}\, dx \to \frac{1}{\delta} \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x \frac{\xi}{\delta}}\, dx$ ie $f(\delta x) \to \delta^{-1} \hat{f}(\delta^{-1}\xi)$

6. If $\mathbf{f}, \mathbf{g} \in \mathcal{S}(\mathbb{R})$ then $\mathbf{f} * \mathbf{g} \in \mathcal{S}(\mathbb{R})$

To prove this we need to show that:

$$\sup_{x\in\mathbb{R}} |x|^k \, |(f * g)^{(k)}(x)| < \infty \quad \forall k, l \geq 0 \tag{30}$$

First start with $l = 0$ to show that:

$$\sup_{x\in\mathbb{R}} |x|^k \, |(f * g)(x)| < \infty \quad \forall k \geq 0 \tag{31}$$

To prove (31) the idea is to break up the domain in a way that uses rapid decrease of the functions. First suppose that $|x| < 2|y|$: Then:

$$|x|^k \, |g(x - y)| < 2^k |y|^k B < 2^k (1 + |y|)^k B \leq A_k (1 + |y|)^k \tag{32}$$

since $g \in \mathcal{S}(\mathbb{R})$ its is bounded by B, say. Note that $A_k = 2^k B$.

The significance of this estimate is that the RHS of (32) is just a polynomial and we know that a Schwartz space function will decay more rapidly than the inverse of any polynomial and we will use that fact shortly.

The second step is to assume that $|x| \geq 2|y|$. Then:

$$|x|^k\,|g(x-y)| \leq |x|^k\,\frac{B'}{|x-y|^k} \leq B'\,\frac{|x|^k}{\frac{|x|^k}{2^k}} = B'\,2^k \tag{33}$$

where we have used the fact that $|x-y| \geq \big||x|-|y|\big| \geq \frac{|x|}{2}$ because $|y| \leq \frac{|x|}{2}$. B' is a constant.

Thus we have:

$$|x|^k\,|g(x-y)| \leq B'\,2^k \leq 2^k\,(1+|y|)^k\,B' = B_k\,(1+|y|)^k \tag{34}$$

If we choose $C_k = \max\{A_k, B_k\}$ we can bound $|x|^k\,|g(x-y)|$ by $C_k\,(1+|y|)^k$.

So going back to (31) we have:

$$\sup_{x\in\mathbb{R}}|x|^k\,|(f*g)(x)| = \sup_{x\in\mathbb{R}}|x|^k\,\int_{-\infty}^{\infty} f(y)\,g(x-y)\,dy\,\Big| \leq C_k\,\int_{-\infty}^{\infty}|f(y)|\,(1+|y|)^k\,dy < \infty \quad \forall k \geq 0 \tag{35}$$

That (35) is true follows from the fact that $|f(y)|$ decreases faster than the inverse of any polynomial in $|y|$ so the integral is bounded for all $k \geq 0$.

Alternatively we can demonstrate the required boundedness in a more detailed way as follows:

$$\begin{aligned}
\Big|\int_{-\infty}^{\infty} f(y)\,g(x-y)\,dy\Big| &\leq \Big|\int_{|y|\leq\frac{|x|}{2}} f(y)\,g(x-y)\,dy + \int_{|y|>\frac{|x|}{2}} f(y)\,g(x-y)\,dy\Big| \\
&\leq \int_{|y|\leq\frac{|x|}{2}} \big|f(y)\big|\,\big|g(x-y)\big|\,dy + \int_{|y|>\frac{|x|}{2}} \big|f(y)\big|\,\big|g(x-y)\big|\,dy \\
&\leq \int_{|y|\leq\frac{|x|}{2}} \big|f(y)\big|\,\frac{a_m}{|x-y|^{m+1}}\,dy \;+\; \int_{|y|>\frac{|x|}{2}} b\,\big|f(y)\big|\,dy \\
&\leq a_m\,\int_{|y|\leq\frac{|x|}{2}} \big|f(y)\big|\frac{2^{m+1}}{|x|^{m+1}}\,dy \;+\; b\int_{|y|>\frac{|x|}{2}} |y|^{-m}\,\underbrace{|y|^m\big|f(y)\big|}\,dy \\
&\leq \frac{a'_m}{|x|^{m+1}}\,\int_{|y|\leq\frac{|x|}{2}} A\,dy + \frac{b2^m}{|x|^m}\int_{|y|>\frac{|x|}{2}} \underbrace{|y|^m\big|f(y)\big|}\,dy \\
&\leq \frac{c_m\,|x|}{|x|^{m+1}} + \frac{b2^m\,B}{|x|^m} \leq \frac{A_m}{|x|^m} + \frac{B_m}{|x|^m} = \frac{C_m}{|x|^m}
\end{aligned} \tag{36}$$

This shows that $(f * g)(x)$ is bounded for all $x \in \mathbb{R}$, for all $m \geq 0$. Thus $f * g \in \mathcal{S}$.

We still have to show that $\sup_{x\in\mathbb{R}} |x|^l \, |(f * g)^{(k)}(x)| < \infty$ for $k > 0$. We first establish that:

$$\Big(\frac{d}{dx}\Big)^k (f * g)(x) = \Big(f * \Big(\frac{d}{dx}\Big)^k\Big)\big(x\big) \; \forall k = 1, 2 \ldots \tag{37}$$

We know that $h(x) = \big(\frac{d}{dx}\big)^k g(x) \in \mathcal{S}(\mathbb{R})$ if $g \in \mathcal{S}(\mathbb{R})$ and so we have something of the form $|x|^l \, |(f * g)(x)|$ where $f, g \in \mathcal{S}(\mathbb{R})$, hence the above steps demonstrate that this product is rapidly decreasing for all $l \geq 0$.

(37) is established by a simple induction. For $k = 1$ for have:

$$\frac{d}{dx}(f * g)(x) = \frac{d}{dx}\int_{-\infty}^{\infty} f(y)\, g(x-y)\, dy = \int_{-\infty}^{\infty} f(y)\, \frac{d}{dx} g(x-y)\, dy = (f * \frac{dg}{dx})(x) \tag{38}$$

Now for the induction step:

$$\begin{aligned}\frac{d^{k+1}}{dx^{k+1}}(f * g)(x) = \frac{d}{dx}\Big(\Big(\frac{d}{dx}\Big)^k (f * g)(x)\Big) &= \frac{d}{dx}\int_{\infty}^{\infty} f(y)\Big(\frac{d}{dx}\Big)^k g(x-y)\, dy \\ &= \int_{\infty}^{\infty} f(y)\Big(\frac{d}{dx}\Big)^{k+1} g(x-y)\, dy = (f * \Big(\frac{d}{dx}\Big)^{k+1})(x)\end{aligned} \tag{39}$$

Differentiation under the integral sign is justified by the rapid decrease of the derivatives of g (see the Appendix for more details).

Tying this all together it follows that $f * g \in \mathcal{S}(\mathbb{R})$.

7. For Schwartz functions, $\mathbf{f * g = g * f}$

$$(f * g)(x) = \int_{-\infty}^{\infty} f(y)\, g(x-y)\, dy = \int_{-\infty}^{\infty} f(x-u)\, g(u)\, du = (g * f)(x) \tag{40}$$

Note that for any Schwartz function F, $\int_{-\infty}^{\infty} F(x)\, dx = \int_{-\infty}^{\infty} F(-x)\, dx$ because the difference $\big|\int_{-N}^{N} F(x)\, dx - \int_{-N}^{N} F(-x)\, dx\big| = \big|\int_{-N}^{N} F(x)\, dx - \int_{N}^{-N} -F(x)\, dx\big| = 0$. Moreover F satisfies translation invariance ie $\forall h \in \mathbb{R}$, $\int_{-\infty}^{\infty} F(x-h)\, dx = \int_{-\infty}^{\infty} F(x)\, dx$. Thus in the above integration where the substitution $u = x - y$ is made, the process is justified since the substitution is a composition of $y \to -y$ and $y \to y - h$ where $h = x$.

8. If $\mathbf{f, g} \in \mathcal{S}(\mathbb{R})$ then $\int_{-\infty}^{\infty} \mathbf{f(x)\,\hat{g}(x)\,dx} = \int_{-\infty}^{\infty} \mathbf{\hat{f}(y)\,g(y)\,dy}$

The proof of this proposition relies upon changing the order of integration for double integrals. Stein and Shakarchi (see [2] page 141) prove this result for the weaker case of moderately decreasing functions and it holds true for Schwartz functions.

9. Fourier inversion: $\mathbf{f} \in \mathcal{S}(\mathbb{R}) \implies \mathbf{f(x)} = \int_{-\infty}^{\infty} \hat{\mathbf{f}}(\xi)\, \mathbf{e^{2\pi i x\xi}}\, \mathbf{d}\xi$

Proving the inversion formula involves the properties of a Gaussian kernel which I will come to, but the structure of the proof is to first show that:

$$f(0) = \int_{-\infty}^{\infty} \hat{f}(\xi)\, d\xi \tag{41}$$

Having established that result if we let $F(y) = f(y + x)$ we then have:

$$f(x) = F(0) = \int_{-\infty}^{\infty} \hat{F}(\xi)\, d\xi = \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{2\pi i x\xi}\, d\xi \tag{42}$$

Noting that $f(y + x) \to \hat{f}(\xi)e^{2\pi i x\xi}$.

The crux of this proof is the assertion in (42) which necessarily involves a claim of convergence of the integral for the class of Schwartz functions. If we take a Gaussian kernel $G_\delta(x) = e^{-\pi\delta x^2}$ then $\hat{G_\delta}(\xi) = \frac{1}{\sqrt{\delta}} e^{\frac{-\pi\xi^2}{\delta}} = K_\delta(\xi)$. Now $K_\delta(\xi)$ is a "good" kernel in the sense that for every $\eta > 0$, $\int_{|x|>\eta} |K_\delta(x)|\, dx \to 0$ as $\delta \to 0$. For more on kernels see [12].

Using property 8 stated above we then have:

$$\int_{-\infty}^{\infty} f(x)\, K_\delta(x)\, dx = \int_{-\infty}^{\infty} \hat{f}(\xi)\, G_\delta(\xi)\, d\xi \tag{43}$$

Since K_δ is a good kernel we have that:

$$\int_{-\infty}^{\infty} f(x)\, K_\delta(x)\, dx \to f(0) \text{ as } \delta \to 0 \tag{44}$$

One way of seeing this is to note that:

$$\Big|\int_{-\infty}^{\infty} f(x)\, K_\delta(x)\, dx\Big| = \Big|\int_{-\infty}^{\infty} f(x)\, \frac{1}{\sqrt{\delta}} e^{\frac{-\pi x^2}{\delta}}\, dx\Big| \leq \int_{-\infty}^{\infty} |f(y\sqrt{\delta})|\, e^{-\pi y^2}\, dy \tag{45}$$

Because of the continuity of f, $f(y\sqrt{\delta}) \to f(0)$ as $\delta \to 0$. Hence we get that the RHS of (45) $\to \int_{-\infty}^{\infty} f(0)\, e^{-\pi y^2}\, dy = f(0)$ since $\int_{-\infty}^{\infty} e^{-\pi y^2}\, dy = 1$. The good kernel is nothing more than a building block of the Dirac delta "function" which picks out the value of the function at $x = 0$.

The continuity property used above can be demonstrated more rigorously. In [[2] pages 133-134] Stein and Shakarchi show that for a moderately decreasing function f:

$$\int_{-\infty}^{\infty} [f(x - h) - f(x)]\, dx \to 0 \text{ as } h \to 0 \tag{46}$$

This can be generalised to $\int_{-\infty}^{\infty} f(x)\, K_\delta(x)\, dx \to f(0)$ as $\delta \to 0$ where f is a Schwartz function. To prove this we need to break the domain of the integral into three pieces: a central "hump" and two symmetrical tails. In the tail pieces the rapid decrease of the Schwartz function is what gets the estimates small enough. In the hump, one needs uniform continuity in order to get the estimate small enough.

We take N large and fixed and $\delta > 0$. We break the integral up as follows:

$$
\begin{aligned}
&\Big| \int_{-\infty}^{\infty} [f(x) - f(0)]\, K_\delta(x)\, dx \Big| \\
&= \Big| \int_{-\infty}^{-N\sqrt{\delta}} [f(x) - f(0)]\, K_\delta(x)\, dx + \int_{-N\sqrt{\delta}}^{N\sqrt{\delta}} [f(x) - f(0)]\, K_\delta(x)\, dx + \int_{N\sqrt{\delta}}^{\infty} [f(x) - f(0)]\, K_\delta(x)\, dx \Big| \\
&\qquad \le \int_{|x| \ge N\sqrt{\delta}} \Big| f(x) - f(0) \Big| K_\delta(x)\, dx + \int_{-N\sqrt{\delta}}^{N\sqrt{\delta}} \Big| f(x) - f(0) \Big|\, K_\delta(x)\, dx
\end{aligned}
\tag{47}
$$

The first integral in (42) represents to the two tails while the other integral is the hump and in it we make the change of varables $y = \frac{x}{\sqrt{\delta}}$. Then:

$$
\int_{-N\sqrt{\delta}}^{N\sqrt{\delta}} \Big| f(x) - f(0) \Big| \frac{1}{\sqrt{\delta}} e^{\frac{-\pi x^2}{\delta}}\, dx = \int_{-N}^{N} \Big| f(y\sqrt{\delta}) - f(0) \Big|\, e^{-\pi y^2}\, dy \le \int_{-N}^{N} \Big| f(y\sqrt{\delta}) - f(0) \Big|\, dy \tag{48}
$$

Since f is a Schwartz function it is uniformly continuous on any closed interval, say $[-N-1, N+1]$ so we can choose δ small enough so that $\sup_{y \in [-N,N]} \Big| f(y\sqrt{\delta}) - f(0) \Big| \le \frac{\epsilon}{4N}$ where $\epsilon > 0$ is arbitrary.

Thus $\int_{-N}^{N} \Big| f(y\sqrt{\delta}) - f(0) \Big|\, dy \le 2N \frac{\epsilon}{4N} = \frac{\epsilon}{2}$. This shows that we can make the hump small enough. Note that although $e^{-\pi y^2}$ on $[-N, N]$ is bounded by $e^{-\pi N^2}$ which is small for large N, we have fixed N and it is not enough to say that $e^{-\pi N^2} \int_{-N}^{N} \Big| f(y\sqrt{\delta}) - f(0) \Big|\, dy$ can be made small because the integral could dominate for some small δ. Uniform continuity ensures that if we fix the interval we can be sure that when the $|y\sqrt{\delta} - 0|$ is small so is $\Big| f(y\sqrt{\delta}) - f(0) \Big|$.

To estimate the tails we use the boundedness of f no matter what y or δ are:

$$
\begin{aligned}
\int_{|x| \ge N\sqrt{\delta}} \Big| f(x) - f(0) \Big| K_\delta(x)\, dx &= \int_{|y| \ge N} \Big| f(y\sqrt{\delta}) - f(0) \Big|\, e^{-\pi y^2}\, dy \le e^{-\pi N^2} \int_{|y| \ge N} \Big| f(y\sqrt{\delta}) - f(0) \Big|\, dy \\
&\le B e^{-\pi N^2} < \epsilon
\end{aligned}
\tag{49}
$$

10. Convolution with a "good" kernel $\mathbf{K}_\delta(\mathbf{x})$:
$\mathbf{f} \in \mathcal{S}(\mathbb{R}) \implies (\mathbf{f} * \mathbf{K}_\delta)(\mathbf{x}) = \int_{-\infty}^{\infty} \mathbf{f}(\mathbf{x} - \mathbf{t})\, \mathbf{K}_\delta(\mathbf{t})\, \mathbf{dt} \to f(x)$ uniformly as $\delta \to 0$

This is a fundamental property and the proof reveals an important interplay between the rapid decrease of the tails of the kernel and Schwartz functions and uniform continuity in order to get the estimate for the integral small enough.

A "good" kernel is represented by a Gaussian of the form:

$$K_\delta(x) = \frac{1}{\sqrt{\delta}} e^{\frac{-\pi x^2}{\delta}} \tag{50}$$

where $\delta > 0$

We know that $\int_{-\infty}^{\infty} K_\delta(x)\, dx = 1$ and we can also show that:

$$\forall \eta > 0, \int_{|x|>\eta} |K_\delta(x)|\, dx \to 0 \text{ as } \delta \to 0 \tag{51}$$

To see why (51) holds we just make te change of variable $u = \frac{x}{\sqrt{\delta}}$ so that the intergal becomes $\int_{|u|>\frac{\eta}{\sqrt{\delta}}} e^{-\pi u^2}\, du$ which approaches 0 as $\delta \to 0$. More precisely, just considering the symmetrical case of the right tail, if we take the upper limit of the integral as $\frac{M}{\sqrt{\delta}}$ where $M > 0$ is at least as big as $\frac{1}{\sqrt{\delta}}$ so that as $M \to \infty$, $\frac{M}{\sqrt{\delta}} \to \infty$ as $\delta \to 0$. By the mean value theorem for the finite interval $[\frac{\eta}{\sqrt{\delta}}, \frac{M}{\sqrt{\delta}}]$, there exists some $\frac{\eta}{\sqrt{\delta}} < y < \frac{M}{\sqrt{\delta}}$ such that $\left| \int_{\frac{\eta}{\sqrt{\delta}}}^{\frac{M}{\sqrt{\delta}}} e^{-\pi u^2}\, du \right| = \left| -2\pi y e^{-\pi y^2} \frac{(M-\eta)}{\sqrt{\delta}} \right| \leq 4\pi y \frac{M}{\sqrt{\delta}} e^{-\pi y^2} \to 0$ as $y \to \infty$ due to the fundamental property of the exponential that it grows faster than any power.

We have to show that $\left|(f * K_\delta)(x) - f(x)\right| = \left| \int_{-\infty}^{\infty} \big(f(x-t) - f(x)\big) K_\delta(t)\, dt \right| \to 0$ uniformly as $\delta \to 0$. To do this we break the integral into three pieces: two tails (where $|t| > \eta$ and one hump (where $|t| \leq \eta$). To estimate the tails we use the rapid decrease of f and the property of $K_\delta(t)$ given in (51). Thus for any $\epsilon > 0$, $\exists \eta > 0$ such that $|f(x)| < \frac{\epsilon}{4}$ for $|x| \geq \eta$

For the hump we use the fact that f is uniformly continuous on any compact (closed, bounded interval) and the fact that $\int_{-\eta}^{\eta} K_\delta(x)\, dx < 1$. Thus, for any $\epsilon > 0$, $\exists \eta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \eta$

Thus we have:

$$
\begin{aligned}
\Big|(f * K_\delta)(x) - f(x)\Big| &= \Big|\int_{-\infty}^{\infty} \Big(f(x-t) - f(x)\Big) K_\delta(t)\, dt\Big| \\
&= \Big|\int_{-\eta}^{\eta} \Big(f(x-t) - f(x)\Big) K_\delta(t)\, dt + \int_{|t|>\eta} \Big(f(x-t) - f(x)\Big) K_\delta(t)\, dt\Big| \\
&\leq \int_{-\eta}^{\eta} \Big|f(x-t) - f(x)\Big| K_\delta(t)\, dt + \int_{|t|>\eta} \Big|f(x-t) - f(x)\Big| K_\delta(t)\, dt \\
&< e^{\frac{-\pi\eta^2}{\delta}} \int_{-\eta}^{\eta} \Big|f(x-t) - f(x)\Big|\, dt + \int_{|t|>\eta} \Big(|f(x-t)| + |f(x)|\Big) K_\delta(t)\, dt \\
&\leq 2\eta\epsilon\, e^{\frac{-\pi\eta^2}{\delta}} + \int_{|t|>\eta} \frac{2\epsilon}{4} K_\delta(t)\, dt < 2\eta\epsilon + \frac{\epsilon}{2} < C\epsilon \ \text{ where } C \text{ is some positive constant}
\end{aligned} \tag{52}
$$

This establishes the uniform convergence as $\delta \to 0$

6 Riemann-Lebesgue Lemma

This is a fundamental result about the decay of the Fourier transform. If $\mathbf{f} \in \mathcal{S}(\mathbb{R})$ then $\lim_{|\xi|\to\infty} \hat{f}(\xi) = 0$. To prove this we need to perform a little fiddle as follows:

$$
\begin{aligned}
\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x\xi}\, dx &= -\int_{-\infty}^{\infty} f(x)\, e^{-2\pi i\xi(x-\frac{1}{2\xi})}\, dx \\
&= -\int_{-\infty}^{\infty} f(y + \frac{1}{2\xi})\, e^{-2\pi i y\xi}\, dy \text{ using the substitution } y = x - \frac{1}{2\xi}
\end{aligned} \tag{53}
$$

Therefore:

$$
\hat{f}(\xi) = -\int_{-\infty}^{\infty} f(x + \frac{1}{2\xi})\, e^{-2\pi i x\xi}\, dx \tag{54}
$$

(changing the dummy variable from y to x)

and

$$
\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x\xi}\, dx \tag{55}
$$

Adding (54) and (55):

$$
\hat{f}(\xi) = \frac{1}{2}\int_{-\infty}^{\infty} [f(x) - f(x + \frac{1}{2\xi})]\, e^{-2\pi i x\xi}\, dx \tag{56}
$$

To investigate the convergence properties of (56) we perform estimates on the tails (where we use the rapid decrease of f) and the hump (where we use uniform continuity). For the tails we know that $\exists B > 0$ such that $|f(x)| \leq \frac{B}{x^2}$ so that we can find an N big enough so that for $|\xi| > 1$ we will have:

$$\Big| \int_{|x|>N} [f(x) - f(x + \frac{1}{2\xi})]\, e^{-2\pi i x\xi}\, dx \Big| \leq \int_{|x|>N} \Big[\, |f(x)| + |f(x + \frac{1}{2\xi})| \Big]\, dx$$
$$\leq \int_{|x|>N} \Big[\, \frac{B}{x^2} + \frac{B}{(x + \frac{1}{2\xi})^2}\, \Big]\, dx < \frac{\epsilon}{2} \quad (57)$$

For the hump where $|x| \leq N$ the function $g_\xi(x) = f(x) - f(x + \frac{1}{2\xi})$ converges uniformly to zero as $|\xi| \to \infty$ since f is uniformly continuous on $[-N, N]$. So, given $\epsilon > 0$, $\exists M > 0$ such that for all $|\xi| > M$ and all $|x| \leq N$, $|f(x) - f(x + \frac{1}{2\xi})| < \frac{\epsilon}{4N}$.

Thus putting it all together:

$$|\hat{f}(\xi)| \leq \int_{|x|>N} \Big[\, |f(x)| + |f(x + \frac{1}{2\xi})| \Big]\, dx + \int_{|x|\leq N} \Big[\, \big|f(x) - f(x + \frac{1}{2\xi})\big| \Big]\, dx$$
$$\leq \frac{\epsilon}{2} + 2N\frac{\epsilon}{4N} = \epsilon \quad (58)$$

This establishes that $\lim_{|\xi|\to\infty} |\hat{f}(\xi)| = 0$

7 If $f \in \mathcal{S}(\mathbb{R})$ then $\hat{f}$ is rapidly decreasing

What has to be shown is that:

$$\sup_{\xi\in\mathbb{R}} |\xi|^n\, |\hat{f}(\xi)| < \infty \text{ for all integers } n \geq 0 \quad (59)$$

We know from the Riemann-Lebesgue lemma that $\hat{f}$ vanishes at infinity but we need to show that as $|\xi| \to \infty$ $\hat{f}$ vanishes faster than the reciprocal of any polynominal. We also know that:

$$f'(x) \xrightarrow{\mathcal{F}} \quad 2\pi i\xi \hat{f}(\xi) \quad (60)$$

(60) leads inductively to this:

$$f^{(n)}(x) \xrightarrow{\mathcal{F}} \quad (2\pi i\xi)^n\, \hat{f}(\xi) \quad (61)$$

That (61) holds follows from integration by parts:

$$\int_{-\infty}^{\infty} f^{(n+1)}(x)\, e^{-2\pi i x\xi}\, dx = \left[f^{(n)}(x)\, e^{-2\pi i x\xi}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f^{(n)}(x)\, (-2\pi i\xi)\, e^{-2\pi i x\xi}\, dx = (2\pi i\xi)^{n+1}\, \hat{f}(\xi) \tag{62}$$

Using the fact that f is a Schwartz function it follows that $\left[f^{(n)}(x)\, e^{-2\pi i x\xi}\right]_{-\infty}^{\infty} = 0$

Similarly, since $f^{(n)}(x) \in \mathcal{S}(\mathbb{R})$ and the Riemann-Lebesgue lemma we see that:

$$\lim_{|\xi|\to\infty} \widehat{f^{(n)}(\xi)} = 0 \tag{63}$$

ie

$$\lim_{|\xi|\to\infty} (2\pi i\xi)^n\, \hat{f}(\xi) = 0 \tag{64}$$

this shows that $\sup_{\xi\in\mathbb{R}} |\xi|^n\, |\hat{f}(\xi)| < \infty$ which means that $\hat{f}(\xi)$ is rapidly decreasing.

8 If $f \in \mathcal{S}(\mathbb{R})$ then $\hat{f} \in \mathcal{S}(\mathbb{R})$

As a first step we note that $\hat{f} \in C^{\infty}$. To establish this we start with the fact that f is a Schwartz function so we have that:

$$\int_{-\infty}^{\infty} (1 + |x|)^k\, |f(x)|\, dx < \infty \tag{65}$$

We also have that $\hat{f}$ is k times differentiable with :

$$\frac{d^k}{d\xi^k} \hat{f}(\xi) = (-2\pi i \widehat{x (f(x)})^k(\xi) \tag{66}$$

To prove (65) we need to establish the base case $k = 1$ which follows from Property 2 (the case of $k = 0$ is simply the case $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$ which is a subset of the comments made below) ie

$$\int_{-\infty}^{\infty} (1 + |x|)\, |f(x)|\, dx < \infty \tag{67}$$

We know that we can make $\int_{|x|>N} |f(x)|\, dx$ small since f is rapidly decreasing. For the hump $\int_{|x|\leq N} |f(x)|\, dx$ we use the fact that f is bounded so the net result is that $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$.

Similarly, for $\int_{|x|>N} |x|\,|f(x)|\,dx$ we use rapid decrease properties to ensure the tails are small enough and for the hump $\int_{|x|\leq N} |x|\,|f(x)|\,dx$ we use the boundedness of f. To finalise the induction we note that:

$$\int_{-\infty}^{\infty} (1+|x|)^{k+1}\,|f(x)|\,dx = \int_{-\infty}^{\infty} (1+|x|)\,(1+|x|)^{k}\,|f(x)|\,dx$$
$$= \int_{-\infty}^{\infty} (1+|x|)^{k}\,|f(x)|\,dx + \int_{-\infty}^{\infty} |x|\,(1+|x|)^{k}\,|f(x)|\,dx = A + B < \infty \quad (68)$$

A is bounded by the induction hypothesis while B is bounded because we can replicate the earlier arguments. Alternatively, since f is a Schwartz function it decays faster than the inverse of any polynomial so that $\int_{-\infty}^{\infty} (1+|x|)^{k}\,|f(x)|\,dx$ is bounded.

9 First proof that $e^{-\pi x^2} \to e^{-\pi \xi^2}$

To prove that the Fourier transform of the Gaussian $f(x) = e^{-\pi x^2}$ is $\hat{f}(\xi) = e^{-\pi \xi^2}$ we may as well use the properties developed above. We start with the definition of the Fourir transform $\hat{f}(\xi) = F(\xi) = \int_{-\infty}^{\infty} e^{-\pi x^2}\, e^{-2\pi i x \xi}\,dx$ so that $F(0) = \int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = 1$. Now property 2 established above ie $-2\pi i x\, f(x) \to \frac{d}{d\xi}\,\hat{f}(\xi)$ allows us to say that:

$$F'(\xi) = \int_{-\infty}^{\infty} f(x)\,(-2\pi i x)\,e^{-2\pi i x \xi}\,dx = i \int_{-\infty}^{\infty} f^{'}(x)\,e^{-2\pi i x \xi}\,dx \quad (69)$$

since $f^{'}(x) = -2\pi x\, f(x)$. We also know from Property 1 ie $f^{'}(x) \to 2\pi i \xi\, \hat{f}(\xi)$ that:

$$F^{'}(\xi) = i(2\pi i \xi)\,\hat{f}(\xi) = -2\pi \xi F(\xi) \quad (70)$$

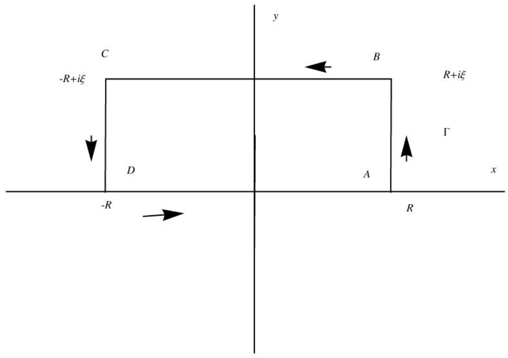
So if $G(\xi) = F(\xi)\,e^{\pi \xi^2}$ then $G^{'}(\xi) = 2\pi \xi F(\xi)\,e^{\pi \xi^2} + e^{\pi \xi^2}\,F^{'}(\xi)$ and hence $G^{'}(0) = 0 + F^{'}(0) = 0$ using (69). This means that $G(\xi)$ is constant for all values of ξ. But since $F(0) = 1$ we must have that G is identically equal to 1 which means that $F(\xi) = \hat{f}(\xi) = e^{-\pi \xi^2}$

10 Second proof that $e^{-\pi x^2} \to e^{-\pi \xi^2}$

This is a classically inspired proof using complex variable theory using Cauchy's Theorem which states that if f is holomorphic or analytic in an open set Ω and $\Gamma \subset \Omega$ is a closed curve whose interior is contained in Ω then:

$$\int_{\Gamma} f(z)\,dz = 0 \quad (71)$$

Now I can hear you saying: "Where does this get us?". The answer is that by suitable choice of $f(z)$ and the contour Γ you can actually get a useful result out of (71). The assumption that f be holomorphic on the relevant open set is easily satisfied in the case of $f(z) = e^{-\pi z^2}$ where z is complex since it is differentiable everywhere in $\mathbb{C}$ (If you can't see why see the Appendix). Thus all we need to do is choose a suitable contour such as the one below so that we can use (71):



Thus we have that:

$$\int_{\Gamma} e^{-\pi z^2}\, dz = 0 \tag{72}$$

and we consider the four paths in the contour as follows.

(i) Along AB $z = R + iy$ and $dz = i\,dy$ where $0 \leq y \leq \xi$. Hence:

$$\Big|\int_{\mathrm{AB}} e^{-\pi z^2}\, dz\Big| = \Big|\int_0^{\xi} e^{-\pi(R^2+2iRy-y^2)}\,(idy)\Big| \leq e^{-\pi R^2} \int_0^{\xi} \Big|e^{-\pi(2iRy-y^2)}\Big|\, dy \leq C\xi\, e^{-\pi R^2} \tag{73}$$

for some constant C noting that $\left|e^{-\pi(2iRy-y^2)}\right| = \left|e^{-2\pi iRy}\right|\left|e^{\pi y^2}\right| \leq e^{\pi\xi^2}$ where ξ is fixed. Hence (73) shows that the integral along AB converges to zeros as $R \to \infty$. Identical logic applies to the path CD where $z = -R + iy$ and $0 \leq y \leq \xi$. Thus $\int_{\mathrm{AB}} e^{-\pi z^2}\, dz = \int_{\mathrm{CD}} e^{-\pi z^2}\, dz = 0$.

Along BC, $z = x + i\xi$ where $-R \leq x \leq R$ and $dz = dx$ (but note that x starts at R to get the right orientation for the path so that the interior is on the inside) hence:

$$\int_{\mathrm{BC}} e^{-\pi z^2}\, dz = \int_{R}^{-R} e^{-\pi(x^2+2ix\xi-\xi^2)}\, dx = -e^{\pi\xi^2} \int_{-R}^{R} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx \tag{74}$$

Finally, along path DA we have $z = x$ for $-R \leq x \leq R$ and $dz = dx$ so that:

$$\int_{\mathrm{DA}} e^{-\pi z^2}\, dz = \int_{-R}^{R} e^{-\pi x^2}\, dx \tag{75}$$

From (72) we have:

$$\int_{\Gamma} e^{-\pi z^2}\, dz = \int_{\mathrm{AB}} e^{-\pi z^2}\, dz + \int_{\mathrm{BC}} e^{-\pi z^2}\, dz + \int_{\mathrm{CD}} e^{-\pi z^2}\, dz + \int_{\mathrm{DA}} e^{-\pi z^2}\, dz = 0 \tag{76}$$

and using (73)-(75) and letting $R \to \infty$ (noting that integral in (75) converges to 1) we have:

$$1 - e^{\pi\xi^2} \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx = 0 \implies \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx = e^{-\pi\xi^2} \tag{77}$$

11 Applying Fourier transform techniques to solve a variant of the heat equation

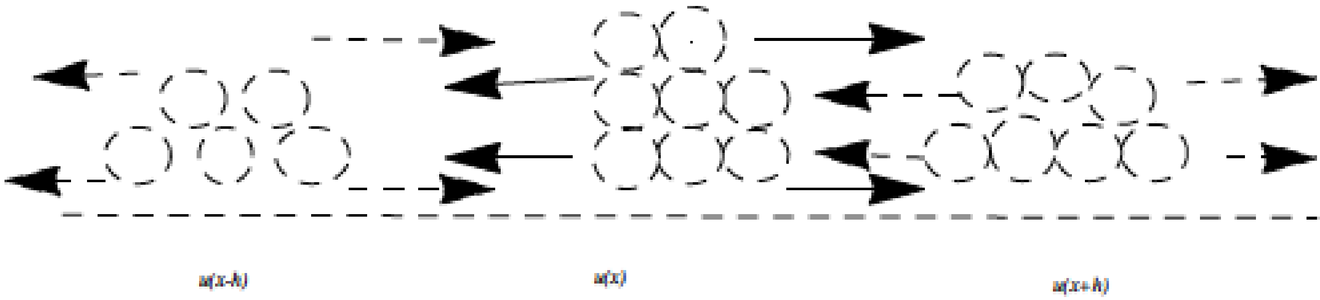
The heat equation is such an important equation in physics it is worth understanding some dimensions to it which are not to my knowledge covered generally in undergraduate engineering courses or indeed courses on financial mathematics. What follows is an expansion of some comments made by the well known partial differential equation expert Luis Cafarelli of the University of Texas, Austin. The heat equation ie $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ represents a diffusion process. A diffusion process such as that represented by the heat equation has a tendency to revert to its surrounding average. To see how this might be the case we need to look at the most simple situation – ie one dimension, which indicates a relationship between diffusion and the Laplacian (in n dimensions the Laplacian is $\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$).

In one dimension the Laplacian of u is simply the second derivative of u and so we look at the limit of the second order incremental quotient. Recall that:

$$u''(x) = \lim_{h\to 0} \frac{\frac{u(x+h)-u(x)}{h} - \left(\frac{u(x)-u(x-h)}{h}\right)}{h} = \lim_{h\to 0} \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} \tag{78}$$

In the diagram below the balls represent particles which can jump left and right in proportion to their number in a pile. The pile at x gains half of the particles coming from adjacent piles and loses its own. This simple rule gives rise to a balance equation of gains (ie $\frac{1}{2}u(x-h) + \frac{1}{2}u(x+h)$) minus losses (ie $u(x)$)which is proportional to:

$$\frac{1}{2}\Big(u(x+h) + u(x-h) - 2u(x)\Big) \tag{79}$$



Equation (79) looks suspiciously like (78) - hence the connection with the Laplacian which has remarkable features: it is rotationally invariant, independent of the system of coordinates and represents a diffusion. As we go up in dimensions we consider the Laplacian as a limit gain-loss of density u at x. We take the average over a unit sphere S of the radial second derivatives in every direction and one of the fundamental results of harmonic analysis is that:

$$\Delta(u) = \fint_S u_{rr}\, dA(s) \tag{80}$$

Recall that for a function u defined in a ball $B(x, r)$ of radius r about x in $\mathbb{R}^n$, with boundary $\partial B(x, r)$ and $\alpha(n)$ is the volume of a unit ball in $\mathbb{R}^n$ and $n\alpha(n)$ is the surface area of the unit ball in $\mathbb{R}^n$, the average of u on $B(x, r)$ is:

$$\fint_{B(x,r)} u(y)\, dy = \frac{1}{\alpha(n) r^n} \int_{B(x,r)} u(y)\, dy \tag{81}$$

In 2 dimensions a function u is harmonic at P (ie it satisfies Laplace's equation $\Delta = 0$) if and only if:

$$u(P) = \frac{1}{2\pi r} \int_{\partial B(P,r)} u\, ds = \frac{1}{\pi r^2} \int_{B(P,r)} u\, dx\, dy \tag{82}$$

To prove (82) we take $P = (x_0, y_0)$ and we suppose that $u(x_0, y_0) = \frac{1}{2\pi r} \int_{\partial B(P,r)} u(x, y)\, ds$ then:

$$\begin{aligned} u(x_0, y_0) = \frac{1}{2\pi r} \int_{\partial B(P,r)} u(x, y)\, ds \; &= \frac{1}{2\pi r} \int_0^{2\pi} u(x_0 + r\cos\theta, y_0 + r\sin\theta)\, r d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} u(x_0 + r\cos\theta, y_0 + r\sin\theta)\, d\theta \end{aligned} \tag{83}$$

The LHS of (83) is simply a constant so if we differentiate with respect to r under the integral sign (and use the chain rule) we get:

$$0 = \frac{1}{2\pi}\int_0^{2\pi}(\frac{\partial u}{\partial x}\cos\theta + \frac{\partial u}{\partial y}\sin\theta)\, d\theta = \frac{1}{2\pi r}\int_0^{2\pi}(\frac{\partial u}{\partial x}\cos\theta + \frac{\partial u}{\partial y}\sin\theta)\, rd\theta$$
$$= \frac{1}{2\pi r}\int_{\partial B(P,r)} \nabla u \,.\, \nu\, ds = \frac{1}{2\pi r}\int_{B(P,r)} div\,(\nabla u)\, dy\, dx = \frac{1}{2\pi r}\int_{B(P,r)} \Delta u\, dy\, dx \quad (84)$$

The divergence theorem justifies the last step in (84). Hence based on our assumption $u(x_0, y_0) = \frac{1}{2\pi r}\int_{\partial B(P,r)} u(x,y)\, ds$ we have shown that $0 = \int_{B(P,r)} \Delta u\, dy\, dx$ for all $r > 0$. If all this holds for every P in some open subset Ω in $\mathbb{R}^2$ then we must have that $\Delta u = 0$ for each such P. Thus u is harmonic.

What this averaging suggests is that the heat equation $\frac{\partial u}{\partial t} = \Delta(u)$ reflects the fact that the density u at the point x makes an infinitesimal comparison within its neighbourhood and tries to revert to the surrounding average.

Problem 1 at pages 169-170 of [2] reads as follows:

The equation:

$$x^2\,\frac{\partial^2 u}{\partial x^2} + ax\,\frac{\partial u}{\partial x} = \frac{\partial u}{\partial t} \quad (85)$$

with $u(x,0) = f(x)$ for $0 < x < \infty$ and $t > 0$ is a variant of the heat equation. This equation can be solved by making the change of variable $x = e^{-y}$ so that $-\infty < y < \infty$. If we set $U(y,t) = u(e^{-y}, t)$ and $F(y) = f(e^{-y})$ then (85) becomes:

$$\frac{\partial^2 U}{\partial y^2} + (1-a)\,\frac{\partial U}{\partial y} = \frac{\partial U}{\partial t} \quad (86)$$

with $U(y,0) = F(y)$. One then has to show that the solution to the original problem is:

$$u(x,t) = \frac{1}{\sqrt{4\pi t}}\int_0^{\infty} e^{\frac{-(\ln(\frac{v}{x})+(1-a)t)^2}{4t}}\, f(v)\frac{dv}{v} \quad (87)$$

We start with:

$$\frac{\partial U}{\partial y} = \frac{\partial u}{\partial x}\,\frac{\partial x}{\partial y} + \frac{\partial u}{\partial t}\,\frac{\partial t}{\partial y} = \frac{\partial u}{\partial x}\,\times(-e^{-y}) = -x\,\frac{\partial u}{\partial x} \quad (88)$$

$$\frac{\partial^2 U}{\partial y^2} = \frac{\partial}{\partial y}(\frac{\partial U}{\partial y}) = \frac{\partial}{\partial x}(-x\frac{\partial u}{\partial x})\,\frac{\partial x}{\partial y} + -x\frac{\partial u}{\partial x}\,\frac{\partial t}{\partial y} = -x(-x\frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x}\times -1)$$
$$= x^2\,\frac{\partial^2 u}{\partial x^2} + x\,\frac{\partial u}{\partial x} \quad (89)$$

$$(1-a)\frac{\partial U}{\partial y} = -x(1-a)\frac{\partial u}{\partial x} \tag{90}$$

$$\frac{\partial U}{\partial t} = \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial t} = \frac{\partial u}{\partial t} \tag{91}$$

Using (89)-(91) we have:

$$\frac{\partial^2 U}{\partial y^2} + (1-a)\frac{\partial U}{\partial y} = x^2\frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial x} - x(1-a)\frac{\partial u}{\partial x} = x^2\frac{\partial^2 u}{\partial x^2} + ax\frac{\partial u}{\partial x} \tag{92}$$

and since $\frac{\partial U}{\partial t} = \frac{\partial u}{\partial t}$ the transformation from (85) to (86) is established.

We now take Fourier transforms of (86) with respect to y as well as the related initial conditions. The Fourier transform of (34) with respect to y is:

$$-4\pi^2\xi^2\hat{U}(\xi,t) + (1-a)2\pi i\xi\,\hat{U}(\xi,t) = \frac{\partial\hat{U}(\xi,t)}{\partial t} \tag{93}$$

Note that the Fourier transform of $\frac{\partial U(y,t)}{\partial t}$ with respect to y is:

$$\int_{-\infty}^{\infty}\frac{\partial U(y,t)}{\partial t}\,e^{-2\pi iy\xi}\,dy = \frac{\partial}{\partial t}\int_{-\infty}^{\infty}U(y,t)\,e^{-2\pi iy\xi}\,dy = \frac{\partial}{\partial t}\hat{U}(\xi,t) \tag{94}$$

See the Appendix for more details about differentiating under the integral sign.

The Fourier transform of the initial condition $U(y,0) = F(y)$ is $\hat{U}(\xi,0) = \hat{F}(\xi)$.

To solve (93) we use the integrating factor $e^{-[-4\pi^2\xi^2+(1-a)2\pi i\xi]t} = e^{-\Phi(\xi)t}$ where $\Phi(\xi) = -4\pi^2\xi^2 + (1-a)2\pi i\xi$. Thus (93) becomes:

$$\Phi(\xi)\,\hat{U}(\xi,t) - \frac{\partial\hat{U}(\xi,t)}{\partial t} = 0 \tag{95}$$

and multiplying by the integrating factor we have:

$$e^{-\Phi(\xi)t}\Big[\Phi(\xi)\,\hat{U}(\xi,t) - \frac{\partial\hat{U}(\xi,t)}{\partial t}\Big] = 0 \tag{96}$$

But (96) is equivalent to:

$$\frac{\partial}{\partial t}\Big\{e^{-\Phi(\xi)t}\;\hat{U}(\xi,t)\Big\} = 0 \tag{97}$$

This means that:

$$e^{-\Phi(\xi)t}\ \hat{U}(\xi,t) = A(\xi) \tag{98}$$

where $A(\xi)$ is some function independent of t.

But $\hat{U}(\xi,0) = \hat{F}(\xi)$ therefore $A(\xi) = \hat{U}(\xi,0) = \hat{F}(\xi)$

So:

$$\hat{U}(\xi,t) = \hat{F}(\xi)\, e^{\Phi(\xi)t} \tag{99}$$

At this point, if we know what function transforms to $e^{\Phi(\xi)t}$ then (99) says that $\hat{U}(\xi,t)$ is the product of two transforms and we can recover the original $U(y,t)$ and hence $u(x,t)$ by convolution ie if $\hat{h}(\xi) = \hat{f}(\xi)\,\hat{g}(\xi)$ then $h(x) = \int_{-\infty}^{\infty} f(x-t)\,g(t)\,dt$

We know that the function which transforms to $e^{\Phi(\xi)t}$ must be a Gaussian with appropriate scaling and the detailed derivation is given below.

We let $\hat{G}(\xi,t) = e^{\Phi(\xi)t} = \int_{-\infty}^{\infty} G(v,t)\, e^{-2\pi i v\xi}\, dv$

Therefore, by Fourier Inversion:

$$G(v,t) = \int_{-\infty}^{\infty} e^{\Phi(\xi)t}\, e^{2\pi i v\xi}\, d\xi = \int_{-\infty}^{\infty} e^{[-4\pi^2\xi^2+(1-a)2\pi i\xi]t}\, e^{2\pi i v\xi}\, d\xi \tag{100}$$

To evaluate (100) we need to complete the square in the exponential factor. Thus we have:

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{[-4\pi^2\xi^2+(1-a)2\pi i\xi]t}\, e^{2\pi i v\xi}\, d\xi &= \int_{-\infty}^{\infty} e^{-4\pi^2 t\xi^2+(1-a)2\pi i\xi t+2\pi i v\xi}\, d\xi = \int_{-\infty}^{\infty} e^{-4\pi^2 t\left\{\xi^2-\frac{2\pi i}{4\pi^2 t}[(1-a)t+v]\xi\right\}}\, d\xi \\
&= \int_{-\infty}^{\infty} e^{-4\pi^2 t\left\{\left[\xi-\frac{i}{4\pi t}[(1-a)t+v]\right]^2+\frac{[(1-a)t+v]^2}{(4\pi t)^2}\right\}}\, d\xi = e^{\frac{-[(1-a)t+v]^2}{4t}} \int_{-\infty}^{\infty} e^{-4\pi^2 t\left\{\xi-\frac{i}{4\pi t}[(1-a)t+v]\right\}^2}\, d\xi \\
&= e^{\frac{-[(1-a)t+v]^2}{4t}} \int_{-\infty-\frac{i}{4\pi t}[(1-a)t+v]}^{\infty+\frac{i}{4\pi t}[(1-a)t+v]} e^{-4\pi^2 t z^2}\, dz
\end{aligned} \tag{101}$$

To evaluate the integral in (101) one can use the contour given in section 10 in the context of showing that the Gaussian is its own Fourier transform and when you reproduce the steps you will find that:

$$e^{\frac{-[(1-a)t+v]^2}{4t}} \int_{-\infty-\frac{i}{4\pi t}[(1-a)t+v]}^{\infty+\frac{i}{4\pi t}[(1-a)t+v]} e^{-4\pi^2 t z^2}\, dz = e^{\frac{-[(1-a)t+v]^2}{4t}}\ \frac{1}{\sqrt{4\pi t}} \tag{102}$$

Another way of "doing" the integral in (101) (apart from using Mathematica or Matlab!) is to view it as basically a real integral with real limits since the imaginary part in the limits contribute a negligible vertical displacement (note that t and v are treated as fixed) so it is effectively this integral (treating x as a real variable): $\int_{-\infty}^{\infty} e^{-4\pi^2 t x^2}\, dx = \frac{1}{\sqrt{4\pi t}}$ since $\int_{-\infty}^{\infty} e^{-\pi x^2}\, dx = 1$. Just think of the crude image of the infinite real integral being rotated a small amount about the origin, Of course the contour integration approach set out in detail in section 5 is the proper way to do the integral and all I am suggesting here is a way of remembering what the result is by way of a useful visualisation.

Going back to (99) we now know what functions transform to $\hat{F}(\xi)$ and $e^{\Phi(\xi)t}$, namely, $F(y)$ (recall that the initial condition was $U(y,0) = F(y)$ and we took the Fourier transform) and $e^{\frac{-[(1-a)t+v]^2}{4t}} \frac{1}{\sqrt{4\pi t}}$. To get back to $U(y,t)$ we take the inverse Fourier transform of (99) which means that we will get the convolution the two functions just mentioned. More specifically we will get:

$$U(y,t) = \int_{-\infty}^{\infty} G(y-v)\, F(v)\, dv = \int_{-\infty}^{\infty} e^{\frac{-[(1-a)t+y-v]^2}{4t}} \frac{1}{\sqrt{4\pi t}} F(v)\, dv \tag{103}$$

Now recalling that $x = e^{-y}$ so that $\ln x = -y$ we also make the substiution $v^* = e^{-v}$ so that $\ln v^* = -v$ and $dv^* = -e^{-v}dv = -v^* dv$ and $F(v) = f(e^{-v}) = f(v^*)$ (go back to the conditions relating to (85) and the integral in (103) becomes:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{\frac{-[(1-a)t+y-v]^2}{4t}} \frac{1}{\sqrt{4\pi t}} F(v)\, dv &= \int_{\infty}^{0} e^{\frac{-[(1-a)t-\ln x+\ln v^*]^2}{4t}} \frac{1}{\sqrt{4\pi t}} \frac{f(v^*)}{-v^*}\, dv^* \\ &= \frac{1}{\sqrt{4\pi t}} \int_{0}^{\infty} e^{\frac{-[(1-a)t+\ln\left(\frac{v^*}{x}\right)]^2}{4t}} \frac{f(v^*)}{v^*}\, dv^* = u(x,t) \end{aligned} \tag{104}$$

Thus we get to the advertised solution given in (87) (the variable v^* replaces v).

12 Using Fourier transform techniques to solve the Black-Scholes equation

The approach to the solution of the problem given in section 11 provides the foundation for solving the Black-Scholes partial differential equation from finance. The original derivation (see [10]) did not explicitly involve Fourier transform techniques. Indeed, the authors merely referred to a well-known undergraduate Fourier series textbook for the solution to what is essentially a heat equation.

Problem 1 at page 170 of [2] sets up the Black-Scholes differential equation so it can be attacked by the Fourier transfrom techniques already covered. The partial differential equation is:

$$\frac{\partial V}{\partial t} + rs\,\frac{\partial V}{\partial s} + \frac{\sigma^2 s^2}{2}\,\frac{\partial^2 V}{\partial s^2} - rV = 0, \quad 0 < t < T \tag{105}$$

This equation is subject to a 'final' boundary condition which is:

$$V(s, T) = F(s) \tag{106}$$

It will be seen that this final condition is essentially an initial condition once we have done an appropriate transformation. With some experience in playing with differential equations such as (105) a reasonable substitution would look like this:

$$V(s, t) = e^{ax+b\tau}\, U(x, \tau) \tag{107}$$

where $x = \ln s$ and $\tau = \frac{\sigma^2}{2}(T - t)$, $a = \frac{1}{2} - \frac{r}{\sigma^2}$ and $b = -(\frac{1}{2} + \frac{r}{\sigma^2})^2$. Although these parameters are given in the problem in what follows they are derived so you can see where they come from. With these substitutions (105) can be reduced to a simple one dimensional heat equation with an initial condition $U(x, 0) = e^{-ax}\, F(e^x)$ which can then be attacked using the Fourier transform techniques explored above. The problem asks you to prove that:

$$V(s, t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T - t)}} \int_0^\infty e^{\frac{-(\log(\frac{s}{s^*})+(r-\frac{\sigma^2}{2})(T-t))^2}{2\sigma^2(T-t)}}\, F(s^*)\, \frac{ds^*}{s^*} \tag{108}$$

In relation to (108) which is problem 2 in [2], page 170 there is a typo - the $\frac{1}{s^*}$ factor is missing. This correction will become clear in the derivation.

Under the substitution $x = \ln s$ and $\tau = \frac{\sigma^2}{2}(T - t)$ we can get expressions for the relevant partial derivatives as follows:

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial s}\,\frac{\partial s}{\partial x} + \frac{\partial V}{\partial \tau}\,\frac{\partial \tau}{\partial x} \tag{109}$$

But $\frac{\partial \tau}{\partial x} = 0$ and $\frac{\partial s}{\partial x} = s$ so (109) becomes:

$$\frac{\partial V}{\partial x} = s\,\frac{\partial V}{\partial s} = \Omega \tag{110}$$

The second partial derivative is:

$$\frac{\partial^2 V}{\partial x^2} = \frac{\partial \Omega}{\partial x} = \frac{\partial \Omega}{\partial s}\,\frac{\partial s}{\partial x} + \frac{\partial \Omega}{\partial \tau}\,\frac{\partial \tau}{\partial x} = \frac{\partial \Omega}{\partial s}\frac{\partial s}{\partial x} = s\frac{\partial}{\partial s}\Big(s\,\frac{\partial V}{\partial s}\Big) = s^2\,\frac{\partial^2 V}{\partial s^2} + s\,\frac{\partial V}{\partial s} = s^2\,\frac{\partial^2 V}{\partial s^2} + \frac{\partial V}{\partial x} \tag{111}$$

Thus we have:

$$s^2 \frac{\partial^2 V}{\partial s^2} = \frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x} \tag{112}$$

We now do the same for the time derivative τ:

$$\frac{\partial V}{\partial \tau} = \frac{\partial V}{\partial s}\frac{\partial s}{\partial \tau} + \frac{\partial V}{\partial t}\frac{\partial t}{\partial \tau} = \frac{\partial V}{\partial t}\frac{\partial t}{\partial \tau} = \frac{-2}{\sigma^2}\frac{\partial V}{\partial t} \tag{113}$$

since $\frac{\partial s}{\partial \tau} = 0$ and $\frac{\partial t}{\partial \tau} = \frac{\partial}{\partial \tau}\{T - \frac{2}{\sigma^2}\tau\} = \frac{-2}{\sigma^2}$

Hence, using ((110)-(113), (105) becomes:

$$-\frac{\sigma^2}{2}\frac{\partial V}{\partial \tau} + r\frac{\partial V}{\partial x} + \frac{\sigma^2}{2}\Big(\frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x}\Big) - rV = 0$$

$$\text{that is } \frac{\sigma^2}{2}\frac{\partial V}{\partial \tau} = \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x^2} + (r - \frac{\sigma^2}{2})\frac{\partial V}{\partial x} - rV \tag{114}$$

Now with the substiution $V = e^{ax+b\tau}\,U(x,\tau)$ the relevant derivatives become:

$$\frac{\partial V}{\partial \tau} = be^{ax+b\tau}U + e^{ax+b\tau}\frac{\partial U}{\partial \tau} = e^{ax+b\tau}\,(bU + \frac{\partial U}{\partial \tau}) \tag{115}$$

$$\frac{\partial V}{\partial x} = e^{ax+b\tau}\frac{\partial U}{\partial x} + ae^{ax+b\tau}\,U = e^{ax+b\tau}\,(aU + \frac{\partial U}{\partial x}) \tag{116}$$

$$\frac{\partial^2 V}{\partial x^2} = e^{ax+b\tau}\Big(\frac{\partial^2 U}{\partial x^2} + a\frac{\partial U}{\partial x}\Big) + ae^{ax+b\tau}\Big(aU + \frac{\partial U}{\partial x}\Big) = e^{ax+b\tau}\Big(\frac{\partial^2 U}{\partial x^2} + 2a\frac{\partial U}{\partial x} + a^2 U\Big) \tag{117}$$

Substituting (115)-(117) into (114) we get:

$$\frac{\sigma^2}{2}\,e^{ax+b\tau}\,(bU+\frac{\partial U}{\partial \tau}) = \frac{\sigma^2}{2}\,e^{ax+b\tau}\Big(\frac{\partial^2 U}{\partial x^2}+2a\frac{\partial U}{\partial x}+a^2U\Big)+(r-\frac{\sigma^2}{2})\,e^{ax+b\tau}\,(aU+\frac{\partial U}{\partial x})-re^{ax+b\tau}U \tag{118}$$

On dividing by the exponential factor we are left with:

$$\frac{\sigma^2}{2}\frac{\partial U}{\partial \tau} = \frac{\sigma^2}{2}\frac{\partial^2 U}{\partial x^2} + (a\sigma^2 + r - \frac{\sigma^2}{2})\frac{\partial U}{\partial x} + [\frac{a^2\sigma^2}{2} + a(r - \frac{\sigma^2}{2}) - r - \frac{b\sigma^2}{2}]\,U \tag{119}$$

To make (119) into a one dimensional heat equation we need:

$$a\sigma^2 + r - \frac{\sigma^2}{2} = 0 \tag{120}$$

and

$$\frac{a^2\sigma^2}{2} + a(r - \frac{\sigma^2}{2}) - r - \frac{b\sigma^2}{2} = 0 \tag{121}$$

(120) shows that $a = \frac{1}{2} - \frac{r}{\sigma^2}$ as given in the statement of the problem. Substituting this value for a into (121) gives:

$$\begin{aligned}
\frac{\sigma^2}{2}(\frac{1}{4} - \frac{r}{\sigma^2} + \frac{r^2}{\sigma^4}) + (\frac{1}{2} - \frac{r}{\sigma^2})(r - \frac{\sigma^2}{2}) - r - \frac{b\sigma^2}{2} = 0 \\
\frac{\sigma^2}{8} - \frac{r}{2} + \frac{r}{2\sigma^2} + \frac{r}{2} - \frac{\sigma^2}{4} - \frac{r^2}{\sigma^2} + \frac{r}{2} - r = \frac{b\sigma^2}{2} \\
\frac{b\sigma^2}{2} = -(\frac{\sigma^2}{8} + \frac{r}{2} + \frac{r^2}{2\sigma^2}) \\
b = -(\frac{1}{4} + \frac{r}{\sigma^2} + \frac{r^2}{\sigma^4}) = -(\frac{1}{2} + \frac{r}{\sigma^2})^2 \quad (122)
\end{aligned}$$

Accordingly, (119) becomes (after dividing through by $\frac{\sigma^2}{2}$):

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2} \tag{123}$$

which is the familiar one dimensional heat equation. The initial condition for this equation is found by putting $\tau = 0$ and hence $T = t$ in (107) giving:

$$V(s, T) = e^{ax}\, U(x, 0) \tag{124}$$

But $V(s, T) = F(s) = F(e^x)$ so the initial condition becomes:

$$U(x, 0) = e^{-ax}\, F(e^x) \tag{125}$$

Taking the Fourier transform of (123) with respect to x we get:

$$\frac{\partial \hat{U}(\xi, \tau)}{\partial \tau} = -4\pi^2\, \xi^2\, \hat{U}(\xi, \tau) \tag{126}$$

Taking the Fourier transform of the initial condition (125) with respect to x we get:

$$\hat{U}(\xi, 0) = \mathcal{F}\{e^{-ax} F(e^x)\} = C(\xi) \tag{127}$$

As before the solution to (126) is:

$$\hat{U}(\xi, \tau) = C(\xi)\, e^{-4\pi^2 \xi^2 \tau} = \hat{U}(\xi, 0)\, e^{-4\pi^2 \xi^2 \tau} \tag{128}$$

Equation (128) is thus the product of two Fourier transforms and so we can get back to $U(x, \tau)$ and hence $V(s, t)$ by convolution of the functions which give rise to those transforms. The Gaussian transform in (128) is easy since we know that a Gaussian transforms to a Gaussian so we only have to get the scaling right. But we know that if $f(x) = e^{-\pi x^2}$ then $\hat{f}(\xi) = e^{-\pi \xi^2}$ and $f(\delta\, x) \to \frac{1}{\delta} \hat{f}(\frac{\xi}{\delta})$ so that by letting $\delta = \frac{1}{\sqrt{4\pi\,\tau}}$ we can see that $e^{-4\pi^2 \xi^2 \tau}$ is the transform of $\frac{1}{\sqrt{4\pi\tau}}\, e^{\frac{-x^2}{4\tau}}$. The other function is even easier to identify since from (125) the function which we want is simply $e^{-ax} F(e^x)$

Thus by convolution we have:

$$U(x, \tau) = \frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{\infty} e^{-ay} F(e^y)\, e^{\frac{-(x-y)^2}{4\tau}}\, dy \tag{129}$$

Now make the folowing substitutions:
$x = \ln s$ for $0 < s < \infty$
$y = \ln s^*$ for $0 < s^* < \infty$ so that $dy = \frac{ds^*}{s^*}$
$\tau = \frac{\sigma^2}{2}(T - t)$ as before
$V(s, t) = e^{ax+b\tau}\, U(x, \tau)$ as before

Using (107) we therefore have:

$$
\begin{aligned}
V(s,t) = e^{ax+b\tau}\,U(x,\tau) &= \frac{e^{a\ln s + b\frac{\sigma^2}{2}(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{-a\ln s^*}\,F(s^*)\,e^{\frac{-[\ln(\frac{s}{s^*})]^2}{2\sigma^2(T-t)}}\,\frac{ds^*}{s^*} \\
&= \frac{e^{b\frac{\sigma^2}{2}(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty F(s^*)\,e^{a\ln(\frac{s}{s^*})}\,e^{\frac{-[\ln(\frac{s}{s^*})]^2}{2\sigma^2(T-t)}}\,\frac{ds^*}{s^*} \\
&= \frac{e^{b\frac{\sigma^2}{2}(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{\frac{-1}{2\sigma^2(T-t)}\left[[\ln(\frac{s}{s^*})]^2 - 2a\sigma^2(T-t)\ln(\frac{s}{s^*})\right]} F(s^*)\frac{ds^*}{s^*} \\
&= \frac{e^{b\frac{\sigma^2}{2}(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{\frac{a^2\sigma^2(T-t)}{2}}\,e^{\frac{-1}{2\sigma^2(T-t)}\left[\ln(\frac{s}{s^*}) - (\frac{1}{2}-\frac{r}{\sigma^2})\sigma^2(T-t)\right]^2} F(s^*)\frac{ds^*}{s^*} \\
&= \frac{e^{\frac{\sigma^2}{2}(T-t)(a^2+b)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{\frac{-1}{2\sigma^2(T-t)}\left[\ln(\frac{s}{s^*}) + (r-\frac{\sigma^2}{2})(T-t)\right]^2} F(s^*)\frac{ds^*}{s^*} \\
&= \frac{e^{\frac{\sigma^2(T-t)}{2}\times\frac{-2r}{\sigma^2}}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{\frac{-1}{2\sigma^2(T-t)}\left[\ln(\frac{s}{s^*}) + (r-\frac{\sigma^2}{2})(T-t)\right]^2} F(s^*)\frac{ds^*}{s^*} \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{\frac{-1}{2\sigma^2(T-t)}\left[\ln(\frac{s}{s^*}) + (r-\frac{\sigma^2}{2})(T-t)\right]^2} F(s^*)\frac{ds^*}{s^*}
\end{aligned}
\tag{130}
$$

recalling that $a^2 + b = (\frac{1}{2} - \frac{r}{\sigma^2})^2 - (\frac{1}{2} + \frac{r}{\sigma^2})^2 = \frac{-2r}{\sigma^2}$

It is not immediately obvious that $V(s,t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_0^\infty e^{\frac{-1}{2\sigma^2(T-t)}\left[\ln(\frac{s}{s^*}) + (r-\frac{\sigma^2}{2})(T-t)\right]^2} F(s^*)\frac{ds^*}{s^*}$ is in the form given in standard options pricing textbooks such as [12, page 180]. For instance, the general form of the Black-Scholes formula for the value of a European call option C struck at K on stock worth S and expiring at time T (assuming no dividends) is:

$$
C = N(d_1)S - e^{-r(T-t)}\,N(d_2)K \tag{131}
$$

where:

$$
N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}}\,dx \tag{132}
$$

$$
d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{133}
$$

$$
d_2 = \frac{\ln(\frac{S}{K}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{134}
$$

To get from (130) to (131) we have to get an expression for $F(s^*)$ and do a change of variables making sure that the limits of integration are appropriately adjusted.

We make the substitution:

$$u = \frac{\ln(\frac{s}{s^*}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \quad \text{so that } -\sigma\sqrt{T-t}\,du = \frac{ds^*}{s^*} \tag{135}$$

The limits of integration $s^* = 0 \rightarrow u = \infty$ and $s^* = \infty \rightarrow u = -\infty$

Note that:

$$s^* = s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} \tag{136}$$

From the statement of the problem we know that $V(s^*, T) = F(s^*)$ and for a European call option $F(s^*) = \max\{s^* - 1, 0\}$. Note here that in terms of the general situation the relevant condition is $\max\{s^* - K, 0\}$ but in the development of this problem $K = 1$ implicitly. This is so because the original change of variables was $x = \ln s$ but in the more general context $s = Ke^x$.

At this stage we have this:

$$V(s,t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_{\infty}^{-\infty} \left\{ e^{\frac{-u^2}{2}}\, F(s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]}) \right\} \times -\sigma\sqrt{T-t}\,du \tag{137}$$

Now $F(s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]}) = \max\{s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} - 1, 0\}$ so that if $F(s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]}) = \max\{s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} - 1\} > 0$ we can write the integral as follows:

$$V(s,t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln s + (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}} e^{\frac{-u^2}{2}} \left\{ s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} - 1 \right\} du \tag{138}$$

For $\max\{s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} - 1\} > 0$ to hold we need $s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} > 1$ or $e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]} > \frac{1}{s}$. Taking logs we have that:

$-[u\sigma\sqrt{T-t} - (r - \frac{\sigma^2}{2})(T-t)] > \ln(\frac{1}{s}) = -\ln s \implies u\sigma\sqrt{T-t} - (r - \frac{\sigma^2}{2})(T-t) < \ln s$ and so $u < \frac{\ln s + (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$. This gives the upper limit for u.

We can now write (138) as two integrals ie $V(s,t) = I - J$ where:

$$I = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln s + (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}} e^{\frac{-u^2}{2}}\, s\,e^{-[u\sigma\sqrt{T-t} - (r-\frac{\sigma^2}{2})(T-t)]}\,du \tag{139}$$

and

$$J = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln s+(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}} e^{\frac{-u^2}{2}}\ du \tag{140}$$

Using the normal distribution function $N(.)$ we can write $J = e^{-r(T-t)} N(d_2)$. Note that this is consistent with (131) with $K = 1$ as has been assumed in this derivation.

To simplify I we have to complete the square and then do another change of variable. The exponential term has exponent:

$$\begin{aligned}\frac{-u^2}{2} - [u\sigma\sqrt{T-t} - (r - \frac{\sigma^2}{2})(T-t)] &= \frac{-[u^2 + 2u\sigma\sqrt{T-t} - 2(r - \frac{\sigma^2}{2})(T-t)]}{2} \\ = \frac{-[(u+\sigma\sqrt{T-t})^2 - \sigma^2(T-t) - 2(r-\frac{\sigma^2}{2})(T-t)]}{2} &= \frac{-(u+\sigma\sqrt{T-t})^2}{2} + r(T-t) \end{aligned}\tag{141}$$

Thus:

$$\begin{aligned} I = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln s+(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}} s\, e^{\frac{-(u+\sigma\sqrt{T-t})^2}{2}}\ e^{r(T-t)}\, du \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln s+(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}} s\, e^{\frac{-(u+\sigma\sqrt{T-t})^2}{2}}\ du \end{aligned}\tag{142}$$

We now make the change of variable:

$$z = u + \sigma\sqrt{T-t} \text{ so that } dz = du \tag{143}$$

The upper limit in I becomes:

$$\begin{aligned} z = \frac{\ln s + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} + \sigma\sqrt{T-t} = \frac{\ln s + (r - \frac{\sigma^2}{2})(T-t) + \sigma^2(T-t)}{\sigma\sqrt{T-t}} \\ = \frac{\ln s + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \end{aligned}\tag{144}$$

So from(142) I becomes:

$$I = \frac{s}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln s + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}} e^{\frac{-z^2}{2}}\, dz = s\, N(d_1) \tag{145}$$

Hence:

$$V(s,t) = sN(d_1) - e^{-r(T-t)}\, N(d_2) \tag{146}$$

which is the same as (131) with $K = 1$. There are, of course, other ways to get to the same result but they also involve many similar error prone steps.

13 References

[1] R Franklin and R Gosling, "The Molecular Structure of Sodium Thymonucleat", Nature, April 25, 1953, pages 740-741

[2] Werner Heisenberg, "The Physical Principles of the Quantum Theory", Dover 1949.

[3] Elias M Stein & Rami Shakarchi, Fourier Analysis: An Introduction, Princeton University Press, 2003

[4] Claude Cohen-Tannoudji, Bernard Diu & Franck Laloë, Quantum Mechanics, Volume 2 , John Wiley & Sons, 1977

[5] Elias M Stein & Rami Shakarchi, Complex Analysis, Princeton University Press, 2003, pages 14-18.

[6] T. W. Körner, "Fourier Analysis", Cambridge University Press, 1988.

[7] Frank Jones, "Lebesgue Integration on Euclidean Space", Bartlett and Jones Publishers, 2001, page 154-55

[8] Murray M Spiegel, "Schaum's Outline Series: Theory and Problems of Advanced Calculus", McGraw Hill, 1963, p.170

[9] Erik Tavila, Necessary and Sufficient Conditions for Differentiating under the Integral Sign, American Mathematcal Monthly, Vol 108, No 6, June-July 2001, pages 544-548, http://www.jstor.org/stable/2695709

[10] Harley Flanders Differentiating under the Integral Sign, American Mathematical Monthly, Vol 80 No 6, June-July 1973, pages 615-627, http://www.jstor.org/stable/2319163

[11] Fischer Black and Myron Scholes,"The Pricing of Options and Corporate Liabilities", The Journal of Political Economy, Vol. 81, No. 3 (May - Jun., 1973), pp. 637-654: http://www.cs.princeton.edu/courses/archive/fall08/cos323/papers/black_scholes73.pdf

[12] Neil A Chriss, Black-Scholes and Beyond - Option Pricing Models,McGraw-Hill, 1997

[13] Peter Haggstrom, "The good, the bad and the ugly of kernels: Why the Dirichlet kernel is not a good kernel ", http://www.gotohaggstrom.com/The%20good,%20the%20bad,%20and%20the%20ugly%20of%20kernels.pdf

[14] John David Jackson, Classical Electrodynamics , Third Edition, John Wiley, 1999

[15] G Russakoff, " A Derivation of the Macroscopic Maxwell Equations", Am. J. Physics. 38, 1188 (1970). This article can be viewed via this link: http://atsol.fis.ucv.cl/dariop/sites/atsol.fis.ucv.cl.dariop/files/A_Derivation_of_the_Macroscopic_Maxwell_Equations_-_G._Russakoff.pdf

[16] Martin Raussen and Christian Skau, "Interview with Abel Prize Recipient Lennart Carleson," Notices of the American Mathematical Society, February 2007, pages 223-229 The interview can be viewed online via this link: http://www.ams.org/notices/200702/comm-carleson.pdf

[17] Ruel V Churchill, "Fourier Series and Boundary Value Problems", Second Edition,McGraw Hill, 1963

[18] T J I'A Bromwich, "An Introduction to the Theory of Infinite Series", Merchant Books 1908 (digitized by Watchmaker Publishing 2008)

[19] W T Gowers, "A New Proof of Szemerédi's Theorem", https://www.dpmms.cam.ac.uk/~wtg10/papers.html

[20] Elias M Stein & Rami Shakarchi, "Functional Analysis: Introduction to Further Topics in Analysis", Princeton University Press, 2011

[21] Jyrki Kauppinen and Jar Partanen, "Fourier Transforms and Spectroscopy", Wiley,2001, page 18

[22] A H Zemanian, "Distribution Theory and Transform Analysis: An Introduction to Generalized Functions with A
Dover Publications, 1965

14 APPENDIX

14.1 Proof that e^{z^2} is differentiable for all $z \in \mathbb{C}$

Using power series techniques one can show that e^z is differentiable for all $z \in \mathbb{C}$ (see [4 pages 14-18]). Using the chain rule which states that if $d : \Omega \to U$ and $g : U \to \mathbb{C}$ are holomorphic then $(g \circ f)^{'}(z) = g^{'}(f(z))\, f^{'}(z)$ for all $z \in \Omega$ we can differentiate e^{z^2} to get $2z\, e^{z^2}$ since z^2 is clearly differentiable (holomorphic) for all $z \in \mathbb{C}$ because the limit $\frac{(z+h)^2-z^2}{h} = 2z + h$ exists everywhere as $h \to 0$.

Alternatively, one can use the Cauchy-Riemann conditions to prove the differentiability on $\mathbb{C}$. Thus if $x = x + iy$ then $e^{z^2} = e^{x^2-y^2} \cos 2xy + i\, e^{x^2-y^2} \sin 2xy = u(x, y) + iv(x, y)$ and a straighforward calculation shows that $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$ and $\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$ where $u(x, y)$ and $v(x, y)$ are continuously differentiable on $\mathbb{C}$. Thus the derivative exists on $\mathbb{C}$ ie the function is holomorphic on $\mathbb{C}$.

14.2 Differentiating under the integral sign

If we assume that $U(y, t)$ lives in Schwartz space it is an easy matter to show that we can differentiate under the integral sign. A basic theorem underpinning differentiation under the integral sign in a Fourier theory context is given in [6, pp.268-9] and runs as follows.

14.2.1 Main Result

Let $g : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ be a continuous function such that $\frac{\partial g(x,t)}{\partial t}$ exists and is continuous. Suppose $\int_{-\infty}^{\infty} |g(x, t)|\, dx$ and $\int_{-\infty}^{\infty} |\frac{\partial g(x,t)}{\partial t}|\, dx$ and exist for each t and that $\int_{|x|>R} \left|\frac{\partial g(x,t)}{\partial t}\right| dx \to 0$ as $R \to \infty$ uniformly in t on every interval $[a, b]$. Then $\int_{-\infty}^{\infty} g(x, t)\, dx$ is differentiable with:

$\frac{d}{dt} \int_{-\infty}^{\infty} g(x, t)\, dx = \int_{-\infty}^{\infty} \frac{\partial g(x,t)}{\partial t}\, dx$

If the function inhabits Schwartz space the continuity, uniform continuity and integral existence assumptions will be satisfied because of the rapid decrease characteristics of the function.

14.2.2 Using Lebesgue integration theory

To prove the Main Result requires a number of steps that involve continuity, uniform continuity, pointwise convergence and the Fundamental Theorem of Calculus. The process is quite intricate and builds on several results that are usually covered in first and second year Analysis courses. There are other approaches based on Lebesgue measure theory which can yield swifter proofs relying basically on the Lebesgue Dominated Convergence Theorem . For instance, in [6] the theorem is stated this way: Assume that for each $x \in J$, $f(x, t)$ is an integrable function of x on X. Assume that for μ a.e. (almost everywhere) $x \in X$, $f(x, t)$ is differentiable at all $t \in J$. Moreover, assume that there exists $h \in L^1(X)$ such that the derivative satisfies for μ a.e. $x \in X$ and all $t \in J$, $\left|\frac{df(x,t)}{dt}\right| \leq h(x)$.

Then F is also differentiable on J and $\frac{dF(t)}{dt} = \int_X \frac{df(x,t)}{dt}\, d\mu(x)$.

The proof runs like this. Let $t \in J$ be fixed and for small $|\delta|$ consider:

$\frac{F(t+\delta)-F(t)}{\delta} = \int_X \frac{f(x,t+\delta)-f(x,t)}{\delta}\, d\mu(x)$

The integrand is continuous at $\delta = 0$: $\lim_{\delta\to 0} \frac{f(x,t+\delta)-f(x,t)}{\delta} = \frac{df(x,t)}{dt}$ exists by hypothesis. The mean value theorem justifies the following statement:

$\left|\frac{f(x,t+\delta)-f(x,t)}{\delta}\right| = \left|\frac{df(x,t^{'})}{dt}\right| \leq h(x)$ for some $t^{'}$ between t and $t+\delta$. Using a theorem which establishes the continuity of the sought after integral (see [7] page 153) we can conclude that $\lim_{\delta\to 0} \frac{F(t+\delta)-F(t)}{\delta} = \int_X \frac{df(x,t)}{dt}\, d\mu(x)$. The Lebesgue Dominated Convergence Theorem ultimately justifies the theorem which established the continuity of the integral.

14.2.3 Not entirely rigorous but still insightful

The general formula for differentiating under the integral sign, which is called Leibnitz's Rule, is proved in some calculus textbooks in a non-rigorous (if you are an analyst) but nevertheless insightful way. One proof [8] goes like this. It may cause some people to fly into a rage because it skates over the need for uniform continuity but it nevertheless gives the "superstructure" for why the process works and a good foundation for how to make it rigorous. So with those qualifications here it is.

Let $\phi(\alpha) = \int_{u_1}^{u_2} f(x,\alpha)\, dx$ where $a \leq \alpha \leq b$ and u_1, u_2 may depend upon α. If $f(x,\alpha)$ and $\frac{\partial f}{\partial \alpha}$ are continuous in both x and α in some region of the x, α plane including $u_1 \leq x \leq u_2$, $a \leq \alpha \leq b$ and if u_1 and u_2 are continuous and have continuous derivatives for $a \leq \alpha \leq b$ then:

$$\frac{d\phi(\alpha)}{d\alpha} = \int_{u_1}^{u_2} \frac{\partial f(x,\alpha)}{\partial \alpha}\, dx + f(u_2,\alpha)\, \frac{du_2}{d\alpha} - f(u_1,\alpha)\, \frac{du_1}{d\alpha} \tag{147}$$

Now if u_1 and u_2 are constants then we get the simple version of the rule: $\frac{d\phi}{d\alpha} = \int_{u_1}^{u_2} \frac{\partial f(x,\alpha)}{\partial \alpha}\, dx$.

To prove this we let $\phi(\alpha) = \int_{u_1(\alpha)}^{u_2(\alpha)} f(x,\alpha)\, dx$. Then:

$$\begin{aligned}
\Delta\phi = \phi(\alpha+\Delta\alpha) - \phi(\alpha) &= \int_{u_1(\alpha+\Delta\alpha)}^{u_2(\alpha+\Delta\alpha)} f(x,\alpha+\Delta\alpha)\, dx - \int_{u_1(\alpha)}^{u_2(\alpha)} f(x,\alpha)\, dx \\
= \int_{u_1(\alpha+\Delta\alpha)}^{u_1(\alpha)} f(x,\alpha+\Delta\alpha)\, dx + \int_{u_1(\alpha)}^{u_2(\alpha)} f(x,\alpha+\Delta\alpha)\, dx &+ \int_{u_2(\alpha)}^{u_2(\alpha+\Delta\alpha)} f(x,\alpha+\Delta\alpha)\, dx - \int_{u_2(\alpha)}^{u_1(\alpha)} f(x,\alpha)\, dx \\
= \int_{u_1(\alpha)}^{u_2(\alpha)} [f(x,\alpha+\Delta\alpha) - f(x,\alpha)]\, dx &+ \int_{u_2(\alpha)}^{u_2(\alpha+\Delta\alpha)} f(x,\alpha+\Delta\alpha)\, dx - \int_{u_1(\alpha)}^{u_1(\alpha+\Delta\alpha)} f(x,\alpha+\Delta\alpha)\, dx
\end{aligned} \tag{148}$$

Now the Mean Value Theorem for integrals ensures that:

$\exists \xi \in (\alpha, \alpha+\Delta\alpha)$
$\exists \xi_1 \in (u_1(\alpha), u_1(\alpha+\Delta\alpha))$
$\exists \xi_2 \in (u_2(\alpha), u_2(\alpha+\Delta\alpha))$

such that:

$\int_{u_2(\alpha)}^{u_1(\alpha)}[f(x,\alpha+\Delta\alpha)-f(x,\alpha)]\,dx = \Delta\alpha \int_{u_2(\alpha)}^{u_1(\alpha)} \frac{\partial f(x,\xi)}{\partial\alpha}\,dx$
$\int_{u_1(\alpha)}^{u_1(\alpha+\Delta\alpha)} f(x,\alpha+\Delta\alpha)\,dx = f(\xi_1,\alpha+\Delta\alpha)[u_1(\alpha+\Delta\alpha)-u_1(\alpha)]$
$\int_{u_2(\alpha)}^{u_2(\alpha+\Delta\alpha)} f(x,\alpha+\Delta\alpha)\,dx = f(\xi_2,\alpha+\Delta\alpha)[u_2(\alpha+\Delta\alpha)-u_2(\alpha)]$

Therefore, $\frac{\Delta\phi}{\Delta\alpha} = \int_{u_1(\alpha)}^{u_2(\alpha)} \frac{\partial f(x,\alpha)}{\partial\alpha}\,dx + f(\xi_2,\alpha+\Delta\alpha)\,\frac{\Delta u_2}{\Delta\alpha} - f(\xi_1,\alpha+\Delta\alpha)\,\frac{\Delta u_1}{\Delta\alpha}$

Finally letting $\Delta\alpha \to 0$ and using continuity of the derivatives we get the result:
$\frac{d\phi(\alpha)}{d\alpha} = \int_{u_1(\alpha)}^{u_2(\alpha)} \frac{\partial f(x,\alpha)}{\partial\alpha}\,dx + f(u_2,\alpha)\,\frac{du_2}{d\alpha} - f(u_1,\alpha)\,\frac{du_1}{d\alpha}$. This final step obscures the role of uniform continuity which is much more prominent in the following line of argument.

14.2.4 Making the Rigour Police happy

To make the Rigour Police happy we have to resort to the quite a bit more machinery. We start with the following theorem:

Theorem1:

Let f_n be a sequence of functions with domain $\mathcal{D}$ uniformly converging to f ie $f_n \rightrightarrows f$. If the f_n are continuous for each n, then f is continuous on $\mathcal{D}$.

We firstly note what uniform continuity means. The sequence of functions f_n converge uniformly to f on a domain $\mathcal{D}$ if for any given $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that , for all $x \in \mathcal{D}$, $|f_n(x) - f(x)| < \epsilon$ whenever $n > N$. Note that this is a much stronger requirement than pointwise convergence: uniform convergence is a global property while pointwise convergence is a local property.

Pointwise convergence amounts to this requirement: If given any $\epsilon > 0$, for each $x \in \mathcal{D}$ there exists a positive integer $N(x)$ (thus N depends on x) such that $|f_n(x) - f(x)| < \epsilon$ whenever $n > N(x)$.

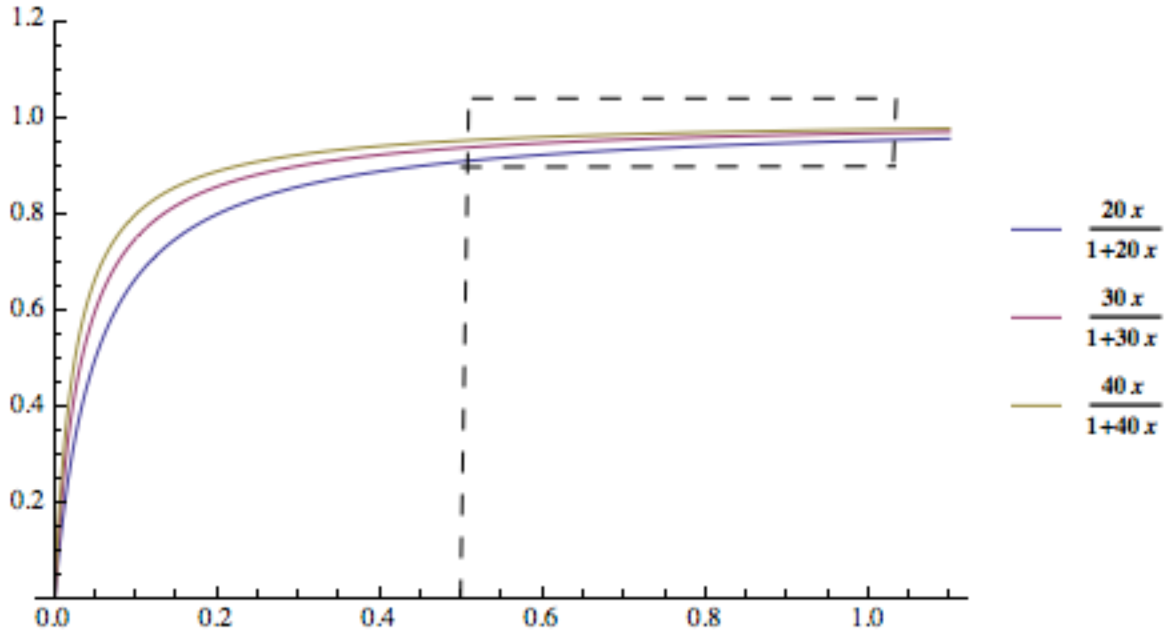
The logical structures of the two definitions are subtly but significantly different. In the case of pointwise convergence $N(x)$ depends on the choice of $x \in \mathcal{D}$ (and on ϵ) while with uniform convergence the N that is determined may depend on ϵ but cannot depend on $x \in \mathcal{D}$. The requirement, "there exists an $N \in \mathbb{N}$ such that , for all $x \in \mathcal{D}$" etc makes this clear.

If we consider the functions defined as follows:

$$g_n(x) = \frac{nx}{nx+1} \qquad \frac{1}{2} \le x \le 1 \tag{149}$$

we can show that the g_n converge pointwise to the limit $g(x) = 1$ for $\frac{1}{2} \le x \le 1$. Thus $|g_n(x) - g(x)| = |\frac{nx}{nx+1} - 1| = \frac{1}{nx+1} \le \frac{1}{nx} < \epsilon$ if we choose $N(x) = 1 + [\frac{1}{\epsilon x}]$. For instance, suppose $\epsilon = 0.1$ and $x = \frac{1}{2}$ then $N(\frac{1}{2}) = 21$ so that $|g_{21}(\frac{1}{2}) - g(\frac{1}{2})| = \frac{1}{\frac{21}{2}+1} = \frac{2}{23} < \frac{1}{10} = \epsilon$.

To show that the $g_n(x)$ converge uniformly to $g(x) = 1$ on $[\frac{1}{2}, 1]$ we note that $|g_n(x) - g(x)| = \frac{1}{nx+1} < \frac{1}{nx} = \frac{\epsilon}{n}\frac{1}{\epsilon x} \leq \frac{\epsilon}{n}\frac{1}{\epsilon}2$ using the fact that $x \geq \frac{1}{2}$ and so $\frac{1}{x} \leq 2$. Thus if we choose $n > \frac{2}{\epsilon}$ (note that this is independent of x) we will have $\frac{1}{nx} < \frac{\epsilon}{2}\frac{1}{x} < \frac{\epsilon}{2}2 = \epsilon$ and so the convergence is uniform on $[\frac{1}{2}, 1]$. See the diagram below:



With this background we can now prove Theorem 1 as follows. Let $x_0 \in \mathcal{D}$ and $\{x_n\}$ be a sequence in $\mathcal{D}$ such that $x_n \to x_0$. We use sequences because we know that f is continuous at x_0 if and only if $f(x_n) \to f(x_0)$ when $x_n \to x_0$. So we choose our $\epsilon > 0$ and note that because $f_n(x) \rightrightarrows f(x)$ there exists a positive integer N_1 such that $|f_n(x_m) - f(x_m)| < \frac{\epsilon}{3}$ for all $m \in N$ and all $n > N_1$. Note here that the requirement $m \in N$ is merely picking up the fact that in the definition of uniform continuity the requirements hold for any $x \in \mathcal{D}$ and so must hold for all elements of the sequence. Moreover, N_1 works for x_0 and all x_m because that is what uniform continuity means.

Because f_n is continuous at x_0 we can find an N such that $|f_n(x_m) - f_n(x_0)| < \frac{\epsilon}{3}$ for all $m > N$

We now do the standard trick of splitting up the relevant absolute difference (ie $|f(x_m) - f(x_0)|$) into, in this case, 3 separate absolute differences which we can separately make arbitrarily small. Thus:

$$
\begin{aligned}
|f(x_m) - f(x_0)| &= |f(x_m) - f_n(x_m) + f_n(x_m) - f_n(x_0) + f_n(x_0) - f(x_0)| \\
&\leq |f(x_m) - f_n(x_m)| + |f_n(x_m) - f_n(x_0)| + |f_n(x_0) - f(x_0)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \quad (150)
\end{aligned}
$$

for all $m > \max\{N_1, N\}$.

This shows that $f(x_m) \to f(x_0)$ which means that f is continuous at x_0 and since x_0 was arbitrary, at all points in $\mathcal{D}$.

We need another theorem as a foundation for the ultimate proof.

Theorem2:

Suppose $\{f_n\}$ is a sequence of functions all of which are integrable on $[a, b]$ and uniformly convergent with limit f. For all $n \in \mathbb{N}$, let $g_n(x) = \int_a^x f_n(t)\, dt$ for $x \in [a, b]$. Then $\{g_n(x)\}$ also converges uniformly on $[a, b]$ and $\lim_{n\to\infty} g_n(x) = g(x)$ where $g(x) = \int_a^x f(t)\, dt$. In other words, we can interchange the limiting process with integration ie $\lim_{n\to\infty} \int_a^x f_n(t)\, dt = \int_x^x \lim_{n\to\infty} f_n(t)\, dt$.

To prove this theorem we take, as usual, $\epsilon > 0$ and note that because the f_n uniformly converge to f there exists an N such that $|f_n(x) - f(x)| < \frac{\epsilon}{b-a}$ for all $x \in [a, b]$ and for all $n > N$. Because we want to show that $g_n(x)$ converges uniformly to $g(x)$ on $[a, b]$ we consider the difference:

$$
\begin{aligned}
|g_n(x) - g(x)| &= |\int_a^x f_n(t)\, dt - \int_a^x f(t)\, dt| = |\int_a^x (f_n(t) - f(t))\, dt| \\
&\leq \int_a^x |f_n(t) - f(t)|\, dt < \frac{\epsilon}{b-a} \int_a^x dt < \frac{\epsilon}{b-a}(b-a) = \epsilon \quad (151)
\end{aligned}
$$

Now because (151) holds for all $x \in [a, b]$ provided $n > N$ where N is independent of x this means that $g_n \rightrightarrows g$ on $[a, b]$.

Theorem3:

Let $\{f_n\}$ be a sequences of functions which are differentiable on $[a, b]$ such that $f_n \to f$ (ie pointwise convergence). If the derivatives f_n' are continuous on $[a, b]$ for all $n \in \mathbb{N}$ and if the sequence f_n' is uniformly convergent, then $\lim_{n\to\infty} f_n' = f'$. Thus the limit of uniformly convergent derivatives is the derivative of the limit.

To prove this theorem we first note that since f_n' is uniformly convergent there must be some function h such that $f_n' \rightrightarrows h$. By Theorem 1 h is continuous on $[a, b]$. Now let $g_n(x) = \int_a^x f_n'(t)\, dt = f_n(x) - f_n(a)$ for $x \in [a, b]$ and $g(x) = \int_a^x h(t)\, dt$ for $x \in [a, b]$. The hypotheses of Theorem 2 are satisfied (note that the f_n' are integrable on $[a, b]$) so that $g_n \to g$. But for every $x \in [a, b]$, $g_n(x) = f_n(x) - f_n(a) \to f(x) - f(a)$ (this is because of the assumed pointwise convergence of f). Hence $g(x) = f(x) - f(a)$ for all x and this limit is unique. By applying the Fundamental Theorem of Calculus we see that since h is continuous, g is differentiable on

$[a, b]$ and $g'(x) = h(x)$. So f is also differentiable on $[a, b]$ and $g' = f' = h$ which finalises the proof.

Theorem4:

Let $g : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ be a continuous function such that $\frac{\partial g(x,t)}{\partial t}$ exists and is continuous. Then $\int_a^b g(x,t)\,dx$ is differentiable with $\frac{d}{dt} \int_a^b g(x,t)\,dx = \int_a^b \frac{\partial g(x,t)}{\partial t}\,dx$

To prove this we need to note that since $\frac{\partial g(x,t)}{\partial t}$ is continuous we can interchange the order of integration in the following integrals [see [3], pages 291-297]:

$$\int_0^y \Big(\int_a^b \frac{\partial g(x,t)}{\partial t}\,dx \Big)\,dt = \int_a^b \Big(\int_0^y \frac{\partial g(x,t)}{\partial t}\,dt \Big)\,dx = \int_a^b \Big(g(x,y) - g(x,0) \Big)\,dx$$
$$= \int_a^b g(x,y)\,dx - \int_a^b g(x,0)\,dx \quad (152)$$

Therefore:

$$\int_a^b g(x,y)\,dx = \int_0^y G(t)\,dt + \text{constant} \quad (153)$$

where $G(t) = \int_a^b \frac{\partial g(x,t)}{\partial t}\,dx$

Recall that if G is continuous at each $y \in \mathbb{R}$ then $\frac{d}{dy}\Big(\int_0^y G(t)\,dt + \text{constant} \Big) = G(y) = \int_a^b \frac{\partial g(x,y)}{\partial y}\,dx$. This flows from the fact that continuity of the partial derivative ensures integrability which in turn ensures continuity of the integral. The Fundamental Theorem of Calculus lets us differentiate the integral.

Because $\frac{\partial g(x,t)}{\partial t}$ is continuous on each closed subset of $\mathbb{R} \times \mathbb{R}$ it is uniformly continuous on all such subsets. This allows us to perform the following estimate:

$$|G(t) - G(y)| = \Big| \int_a^b \frac{\partial g(x,t)}{\partial t}\,dx - \int_a^b \frac{\partial g(x,y)}{\partial y}\,dx \Big| = \Big| \int_a^b \Big(\frac{\partial g(x,t)}{\partial t} - \frac{\partial g(x,y)}{\partial y} \Big)\,dx \Big|$$
$$\leq \int_a^b \Big| \frac{\partial g(x,t)}{\partial t} - \frac{\partial g(x,y)}{\partial y} \Big|\,dx \leq (b-a)\,\epsilon \quad (154)$$

where we have used the fact that the partial derivative is uniformly continuous.

This establishes that G is continuous and so $\int_a^b g(x,y)\,dx$ is differentiable. Thus:

$$\frac{d}{dy} \int_a^b g(x,y)\,dx = G(y) = \int_a^b \frac{\partial g(x,y)}{\partial y}\,dx \quad (155)$$

14.2.5 Proof of the Main Result

Let $f_n(t) = \int_{-n}^{n} g(x,t)\,dx$ and $f(t) = \int_{-\infty}^{\infty} g(x,t)\,dx$. Then:

$|f_n(t) - f(t)| = \left| \int_{-n}^{n} g(x,t)\,dx - \int_{-\infty}^{\infty} g(x,t)\,dx \right| = \left| \int_{|x|>n} g(x,t)\,dx \right|$
$\leq \int_{|x|>n} \left| g(x,t) \right| dx \to 0$ since by assumption $\int_{-\infty}^{\infty} |g(x,t)|\,dx$ exists and so the tails must approach 0 for sufficiently large n. Thus $f_n \to f$ pointwise.

Now f_n is differentiable for all n where $f_n'(t) = \frac{d}{dt} \int_{-n}^{n} g(x,t)\,dx = \int_{-n}^{n} \frac{\partial g(x,t)}{\partial t}\,dx$ using Theorem 4. Furthermore $\left| f_n'(t) - \int_{-\infty}^{\infty} \frac{\partial g(x,t)}{\partial t}\,dx \right| = \left| \int_{-n}^{n} \frac{\partial g(x,t)}{\partial t}\,dx - \int_{-\infty}^{\infty} \frac{\partial g(x,t)}{\partial t}\,dx \right|$
$\leq \int_{|x|>n} \left| \frac{\partial g(x,t)}{\partial t} \right| dx \to 0$ uniformly on each interval $[a,b]$ by assumption.

Thus $f(t) = \int_{-\infty}^{\infty} g(x,t)\,dx$ is differentiable and $\frac{d}{dt} \int_{-\infty}^{\infty} g(x,t)\,dx = \int_{-\infty}^{\infty} \frac{\partial g(x,t)}{\partial t}\,dx$

For more reading on differentiating under the integral sign see [9] and [10].

15 Proof that $\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$

This integral was pivotal to Dirichlet's work on Fourier series in the 19^{th} century. There are various ways to prove it and Dirichlet's original proof is referred to in [18] at page 446. Churchill's proof ([17] pages 85-86) is relatively straightforward and runs as follows.

We first note some basic properties of $S(x) = \frac{\sin x}{x}$ for $x \neq 0$ and $S(0) = 1$.

$$S(x) = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \dots \tag{156}$$

$S(x)$ is even and the series in (156) is an alternating one. Because we can write $S(x) = 1 - (\frac{x^2}{3!} - \frac{x^4}{5!}) - (\frac{x^6}{7!} - \frac{x^8}{9!}) + \dots$, for $0 < x \leq 1$ it follows that $0 < S(x) < 1$. When $x > 1$ then $|S(x)| < |\sin x|$ so we can say that $|S(x)| \leq 1$ for all $x \geq 0$.

Now to get an estimate for the integral we simply break the domain of integration up into intervals of length π as follows: $(0,\pi), (\pi, 2\pi), \dots$ and note that the sign of $S(x)$ alternates on each successive interval. Thus if we choose $v > 0$ and n as the greatest integer such that $n\pi \leq v$ we will have:

$$\int_0^v S(x)\,dx = \int_0^\pi S(x)\,dx + \int_\pi^{2\pi} S(x)\,dx + \dots + \int_{(n-1)\pi}^{n\pi} S(x)\,dx + \int_{n\pi}^{v} S(x)\,dx \tag{157}$$

Now if we let $A_k = \left| \int_{k\pi}^{(k+1)\pi} S(x)\,dx \right|$ we can write (157) in the following form:

$$\int_0^v S(x)\,dx = A_0 - A_1 + A_2 - \cdots + (-1)^{n-1} A_{n-1} + (-1)^n \theta_n A_n \tag{158}$$

Note that $\theta_n \in [0,1)$. Now $A_0 < \pi$ because $A_0 = |\int_0^\pi S(x)\,dx| \leq \int_0^\pi |S(x)|\,dx < \pi$. On the k^{th} interval ($k > 0$) we have that $|S(x)| < \frac{1}{k\pi}$ so that $A_k < \pi \frac{1}{k\pi} = \frac{1}{k}$. We thus see that $A_{k+1} < A_k$. Because the A_k constitute a sequence of positive terms which converge to zero as $k \to \infty$ the alternating series $\lim_{n\to\infty} \sum_{k=0}^{n-1} (-1)^k A_k$ converges.

The improper integral in (158) exists as $v \to \infty$ since $\theta_n A_n \to 0$ as $n \to \infty$.

If we write:

$$F(t) = \int_0^\infty e^{-tx} \frac{\sin x}{x}\,dx \text{ for } t \geq 0 \tag{159}$$

we can be sure that the integral in (159) converges because $|F(t)| \leq \int_0^\infty e^{-tx} |\frac{\sin x}{x}|\,dx \leq \int_0^\infty e^{-tx}\,dx = \frac{1}{t}$

We can write (158) as follows because the integral exists for $t \geq 0$:

$$F(t) = \lim_{n\to\infty} \sum_{k=0}^{n-1} (-1)^k \int_{k\pi}^{(k+1)\pi} e^{-tx} \frac{\sin x}{x}\,dx \tag{160}$$

We know from the discussion following (145) that if we take the absolute value of $F(t)$ in (160) we will have for every n, $|F(t)| \leq \sum_{k=0}^{n-1} \int_{k\pi}^{(k+1)\pi} \left|\frac{\sin x}{x}\right| dx$, and the absolute value of the remainder term $A_n \leq \frac{1}{n}$ which is independent of t. This gives us uniform continuity of the series and hence continuity of $F(t)$. Thus $F(+0) = F(0) = \int_0^\infty \frac{\sin x}{x}\,dx$.

We can get an expression for $F^{'}(t)$ by differentiating under the integral sign (which is kosher - see the earlier discussion in the Appendix) and then integrating by parts.

Thus:

$$F^{'}(t) = -\int_0^\infty e^{-tx} \sin x\,dx \tag{161}$$

Now integrating (161) by parts we get:

$$\begin{aligned} -\int_0^\infty e^{-tx} \sin x\,dx &= \left[e^{-tx}\cos x\right]_0^\infty + t\int_0^\infty e^{-tx} \cos x\,dx = -1 + t\int_0^\infty e^{-tx}\cos x\,dx \\ &= -1 + t\left\{\left[e^{-tx}\sin x\right]_0^\infty + t\int_0^\infty e^{-tx}\sin x\,dx\right\} = -1 + t^2 \int_0^\infty e^{-tx}\sin x\,dx \end{aligned} \tag{162}$$

Therefore:

$$(1+t^2)\int_0^{\infty} e^{-tx}\,\sin x\,dx = -1 \implies F^{'}(t) = \frac{-1}{1+t^2} \tag{163}$$

From (163) it follows that:

$$F(t) = -\arctan t + C \tag{164}$$

Because $|F(t)| \leq \frac{1}{t}$ it follows that $F(t) \to 0$ as $t \to \infty$. Thus $C = \frac{\pi}{2}$ because $\arctan t \to \frac{\pi}{2}$ as $t \to \infty$. Finally, because $F(0) = \int_0^{\infty} \frac{\sin x}{x}\,dx$ we have the result: $\int_0^{\infty} \frac{\sin x}{x}\,dx = \frac{\pi}{2}$.