A Collection of Exercises From Real Analysis II

Owen Drummond
Department of Mathematics, Rutgers University
owen.drummond@rutgers.edu

1 Introduction

The following exercises are an assortment of homework problems from *MATH 502 - Theory of Functions of a Real Variable II* with Professor Dennis Kriventsov at Rutgers University. Topics include Radon Measures, weak convergence, Haar measures, Fourier analysis, PDEs, distributions, and probability theory.

Problem 1 (Folland 7.24). *Find examples of sequences $\{\mu_n\}$ in $M(\mathbb{R})$ such that*

(a) $\mu_n \to 0$ vaguely, but $\|\mu_n\| \not\to 0$.

(b) $\mu_n \to 0$ vaguely, but $\int f\, d\mu_n \not\to \int f\, d\mu$ for some bounded measurable f with compact support.

(c) $\mu_n \geq 0$ and $\mu_n \to 0$ vaguely, but there exists $x \in \mathbb{R}$ such that $F_n(x) \not\to F(x)$ (notation as in Proposition 7.19).

Proof. (a) Let $\mu_n = \delta_n$. First, $\mu_n \to 0$ since $\int_{\mathbb{R}} f d\mu_n = \int_{\mathbb{R}} f d\delta_n = f(n) = 0$ as $n \to 0$ since f vanishes at ∞, yet $||\mu_n|| = \delta_n(\mathbb{R}) = 1$, and thus $||\mu_n|| \nrightarrow 0$.

(b) Let $\mu_n = \delta_{1/n} - \delta_{-1/n}$. $\mu_n \to 0$ since $\int_{\mathbb{R}} f d[\delta_{1/n} - \delta_{-1/n}] = f(1/n) - f(-1/n) \to 0$ as $n \to \infty$. Now let $f = \chi_{[0,1]}$. It is easy to see that f is bounded and compactly supported, yet $\int_{\mathbb{R}} f d\mu_n = f(1/n) - f(-1/n) = 1 \neq 0$ for all $n \in \mathbb{N}$ and thus $\int_R f d\mu_n \nrightarrow \int_R f d\mu = 0$.

(c) Let $\mu_n = \delta_{-n}$. We have that $\mu_n \to 0$ since $\int_{\mathbb{R}} f d\delta_{-n} = f(-n) \to 0$ as $n \to \infty$. Now for any $x \in \mathbb{R}, \exists n \in \mathbb{N}$ s.t. $n > x \implies F_n(x) = \mu_n((-\infty, x]) = \delta_{-n}((-\infty, x]) = 1$ as $-n < -|x| < n \implies F_n(x) \nrightarrow F(x) = 0$

□

Problem 2 (Folland 7.27). *Let $C^k([0,1])$ be as in Exercise 9 in §5.1. If $I \in C^k([0,1])^*$, there exist $\mu \in M([0,1])$ and constants $c_0, \ldots, c_{k-1}$, all unique, such that*

$$I(f) = \int f^{(k)}\, d\mu + \sum_{j=0}^{k-1} c_j f^{(j)}(0).$$

(The functionals $f \mapsto f^{(j)}(0)$ could be replaced by any set of k functionals that separate points in the space of polynomials of degree $< k$.)

Proof. Given that $f \in C^k$, using Taylor's Theorem, we can write $f(x) = T_{k-1}(x) + R_k(x)$, where

$$T_{k-1}(x) = f(0) + f'(0)x + \cdots + \frac{f^{(k-1)}(0)}{(k-1)!}x^{k-1}$$

and

$$R_k(x) = \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} f^{(k)}(t)dt$$

and by the linearity of I, we have that $I(f(x)) = I(T_{k-1}(x)) + I(R_k(x))$. By the Riesz Representation Theorem, $I(f) = \int_0^1 f d\tilde{\mu}$, for a unique Radon measure $\tilde{\mu}$. Thus,

$$\begin{aligned} I(T_{k-1}(x)) &= I(f(0)) + I(f'(0)x) + \cdots + I(\frac{f^{(k-1)}(0)}{(k-1)!}x^{k-1}) \\ &= \int_0^1 f(0)d\tilde{\mu} + f'(0)\int_0^1 x d\mu + \cdots + \frac{f^{(k-1)}(0)}{(k-1)!}\int_0^1 x^{k-1}d\tilde{\mu} \\ &= c_0 + f'(0)c'_1 + \ldots + c'_{k-1} \end{aligned}$$

Now each c'_i is uniquely determined by $\tilde{\mu}$ as basis elements of $C^*([0,1])$. Setting $c_i = \frac{c'_i}{(k-1)!}$, we obtain the desired sum, namely $I(T_{k-1}(x)) = \sum_{j=1}^{k-1} c_j f^{(j)}(0)$. Next,

$$\begin{aligned} I(R_k(x)) &= \int_0^1 \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} f^{(k)}(t)dtd\tilde{\mu}(x) \\ &= \int_0^1 \int_0^1 \chi_{[0,x]} \frac{(x-t)^{k-1}}{(k-1)!} f^{(k)}(t)dtd\tilde{\mu}(x) \\ &= \int_0^1 f^{(k)}(t)[\int_0^1 \chi_{[0,x]} \frac{(x-t)^{k-1}}{(k-1)!} d\tilde{\mu}(x)]dt \end{aligned}$$

by Fubini's theorem since both $f^{(k)}$ and $\frac{(x-t)^{k-1}}{(k-1)!}$ are functions in $C([0,1])$. Lastly, setting $g(t) = \int_0^1 \chi_{[0,x]} \frac{(x-t)^{k-1}}{(k-1)!} d\tilde{\mu}(x)$, we have the integral $\int_0^1 f^{(k)}(t)g(t)dt$, and by setting $d\mu = g(t)dt$, we obtain that $I(R_k(x)) = \int_0^1 f^{(k)}d\mu$, as required

$$\implies I(f) = I(T_{k-1}(x)) + I(R_k(x)) = \int f^{(k)}d\mu + \sum_{j=1}^{k-1} c_j f^{(j)}(0)$$

□

Problem 3. *Let $\mu \in M(\mathbb{R}^n)$ be a signed Radon Measure with $\int f d\mu$ for all $f \in C_c^\infty(U)$ for some open set U. Show that $|\mu(U)| = 0$.*

Proof. Assume for the sake of a contradiction that $|\mu(U)| > 0$. Then either $\mu^+(W) > 0$ or $\mu^-(W) > 0$ for some $W \subset K \subset U$, where K is compact. We can construct a sequence of smooth functions approximating χ_W, and $\text{supp}(\chi_W) \subset K$. Denote $\{f_n\}_{n\in\mathbb{N}}$ as this sequence, s.t.

$$\int_{A\subset\mathbb{R}^n} f_n(x)d\mu(x) = \mu(A)$$

and thus

$$\int_U f_n d\mu \to \mu(U) \geq \mu(W) = \mu^+(W) - \mu^-(W) > 0$$

which is a contradiction, and therefore $|\mu(U)| = 0$ $\square$

Problem 4. *(a) Prove a kind of converse to Proposition 0.1 above: assume that μ_k, μ are (positive) Radon measures on X with $\sup_k \|\mu_k\| < \infty$,*

(a) for all open U

$$\mu(U) \leq \liminf_{k\to\infty} \mu_k(U)$$

(b) for all compact K.

$$\mu(K) \geq \limsup_{k\to\infty} \mu_k(K)$$

Show that $\mu_k \overset{*}{\rightharpoonup} \mu$. [Hint: given $f \in C_0(X)$ nonnegative, write

$$\int f\, d\mu_k = \int_0^\infty \mu_k(\{f > t\})\, dt,$$

show that the integrands converge for all but countably many t, and use dominated convergence.]

(b) Assume for $\mu_k \in M(X)$, $\mu_k \overset{*}{\rightharpoonup} \mu \in M(X)$ and $\|\mu_k\| \to \|\mu\|$. Show that $|\mu_k| \overset{*}{\rightharpoonup} |\mu|$. [Hint: To check (0.2), use that $X \setminus K$ is open and Proposition 0.1.]

Proof. (a) Write $\int f d\mu_k = \int_0^\infty \mu_k(\{f > t\})dt$ for a given $f \in C_0(X)$. Note that for some large N, $\exists K_N$ compact such that $\forall x \notin K_N$, $f(x) < \epsilon$ for any $\epsilon > 0$, so for $t \geq \epsilon$, we have that $\{f > t\} \subset K_N$, and hence $\mu_k(\{f > t\}) \leq \mu_k(K_N) < \infty$ since the μ_k's are Radon and therefore finite on compact sets. Moreover, for any $\epsilon > 0 \exists K$ compact such that $K \subset X$ and $\mu_k(X \setminus K) < \epsilon$, and so the μ_k's are uniformly bounded. Now note that $\{f > t\}$ is an open set since f is continuous, and hence

$$\mu(\{f > t\}) \leq \liminf \mu_k(\{f > t\})$$

However, this set $\{f > t\}$ can also be approximated by some compact set such that the difference is measure zero, i.e.

$$\limsup \mu_k(\{f > t\}) \leq \mu(\{f > t\}) \leq \liminf \mu_k(\{f > t\})$$

which demonstrates that indeed $\mu_k(\{f > t\}) \to \mu(\{f > t\})$ expect on a set of Lebesgue Measure zero, so for countably many t. Lastly, $\forall t \in (0, \infty)$, we have that $\mu_k(\{f > t\}) \leq \sup_k \mu_k(\{f > t\}) \leq \sup_k ||\mu_k|| < \infty$. So $g(t) = \sup_k \mu_k(\{f > t\})$ is the function that dominates $\mu_k(\{f > t\})$. Thus, we can now employ dominated convergence:

$$\lim_{k\to\infty} \int_0^\infty \mu_k(\{f > t\})dt = \int_0^\infty \lim_{k\to\infty} \mu_k(\{f > t\})dt = \int_0^\infty \mu(\{f > t\})dt = \int f d\mu$$

and so $\mu_k \overset{*}{\rightharpoonup} \mu$ as required.

(b) Given that $||\mu_k|| \to ||\mu||$, the measures are converging in total variation, and thus $|\mu_k| \to |\mu|$. Moreover, $|\mu|(X)$ is finite:

$$\exists K \text{compact}, |\mu|(X \setminus K) \leq \liminf |\mu_k|(x \setminus K) < \epsilon, \forall \epsilon > 0$$

by Proposition 0.1. Combined with the fact that $\mu_k \overset{*}{\rightharpoonup} \mu$, the $|\mu_k|$'s are uniformly bounded. Without loss of generality, assume f is nonnegative, $|fd|\mu_k|| \leq (f+1)\sup_k ||\mu_k|| = g_k(x) \forall k$. Thus, we can apply dominated convergence to conclude:

$$\lim_{k\to\infty} \int_X fd|\mu_k| = \int_X \lim_{k\to\infty} fd|\mu_k| = \int_X fd|\mu|$$

Thus, $|\mu_k| \overset{*}{\rightharpoonup} |\mu|$

□

Problem 5. *Let $\phi : \mathbb{R}^n \to [0,1]$ be a smooth radial function with $\int \phi = 1$, and set $\phi_r(x) = \frac{1}{r^n}\phi\left(\frac{x}{r}\right)$. Let μ be a signed Radon measure on $\mathbb{R}^n$, and*

$$\mu_r(E) = \iint_{E\times\mathbb{R}^n} \phi_r(x-y)\, d\mu(y)\, dx.$$

Show that as $r \to 0$, $\mu_r \overset{*}{\rightharpoonup} \mu$ and $|\mu_r| \overset{*}{\rightharpoonup} |\mu|$ (i.e. in the weak- topology). Do they converge in $M(\mathbb{R}^n)$ norm topology?*

Proof. As the measure μ_r is the convolution of μ with the smooth radial function ϕ_r. $\phi_r(x) = \frac{1}{r^n}\phi(\frac{x}{r})$ is essentially a rescaled version of ϕ, and since $\int \phi = 1$, ϕ acts as an approximation of the identity function. Moreover, as $r \to 0$, $\phi_r(x)$ approaches the Dirac delta function of x. This, $\mu_r = \phi_r * \mu$ should converge to μ as $r \to 0$. Thus, we have

$$\int_{\mathbb{R}^n} fd\mu_k = \int_{\mathbb{R}^n} f(\phi_r * \mu) = \iint_{\mathbb{R}^n\times\mathbb{R}^n} f(x)\phi_r(x-y)d\mu(y)dx$$

Now since ϕ, and by extension the ϕ_k's are smooth and hence bounded, and f is bounded since it is continuous with compact support, we can bound $f(x)\phi_r(x-y)$ by some function $M(y)$, which is integrable w.r.t $\mu(y)$. Therefore, by an application of Fubini's Theorem and Dominated convergence, we have that

$$\begin{aligned}
\lim_{r\to 0}\int_{\mathbb{R}^n} fd\mu_k &= \lim_{r\to 0}\iint_{\mathbb{R}^n\times\mathbb{R}^n} f(x)\phi_r(x-y)d\mu(y)dx \\
&= \lim_{r\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)\phi_r(x-y)dxd\mu(y) \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \lim_{r\to 0} f(x)\phi_r(x-y)dxd\mu(y) = \int_{\mathbb{R}^n} f(y)d\mu(y) = \int_{\mathbb{R}^n} fd\mu
\end{aligned}$$

and so $\mu_r \overset{*}{\rightharpoonup} \mu$ as desired. The fact that $|\mu_r| \overset{*}{\rightharpoonup} |\mu|$ is a consequence of the fact that $||\mu_r|| \to ||\mu||$. However, in the norm topology, this is not true in the norm topology $M(\mathbb{R}^n)$, as we do not have the same uniform convergence of $\phi_r(x)$ to δ_x, and thus the limit does not converge within the integral. □

Problem 6 (Folland 11.2)**.** *If μ is a Radon measure on the locally compact group G and $f \in C_c(G)$, the functions $x \mapsto \int L_x f\, d\mu$ and $x \mapsto \int R_x f\, d\mu$ are continuous.*

Proof. First, fix $g \in G$ and $\epsilon > 0$. We define the compact subsets K_0, K_1 in the following way: $K_0 = \text{supp} f$, and K_1 as a compact neighborhood of g. Define $K = K_0 K_1^{-1}$. Since $f \in C_c(G)$, f is left-uniform continuous, and hence we have that $||f - L_v(f)||_u < \frac{\epsilon}{\mu(k)+1}$ for any $v \in V$ since $\mu(K)$ is finite, and $U = K_1 \cap gV$, so now we have that

$$\begin{aligned} |\int_G L_g f d\mu - \int_G L_u f d\mu| &\leq |\int_G |L_g f - L_u f| d\mu \\ &\leq \int_K |L_g f - L_u f| d\mu \\ &\leq \int_K |L_g f - L_g L_{g^{-1}u} f| d\mu \\ &\leq \mu(K) ||f - L_{g^{-1}u} f||_U < \epsilon \end{aligned}$$

The same argument can be repeated to show $x \mapsto \int R_x f\, d\mu$ is also a continuous map. □

Problem 7 (Folland 11.3)**.** *Let G be a locally compact group that is homeomorphic to an open subset U of $\mathbb{R}^n$ in such a way that, if we identify G with U, left translation is an affine map — that is, $xy = A_x(y) + b_x$ where A_x is a linear transformation of $\mathbb{R}^n$ and $b_x \in \mathbb{R}^n$. Then $|\det A_x|^{-1}\, dx$ is a left Haar measure on G, where dx denotes Lebesgue measure on $\mathbb{R}^n$. (Similarly for right translations and right Haar measures.)*

Proof. Define $m(E) = \int_E |det A_x|^{-1} dx$ for $E \subset G$. We want to show for a Borel $E \subset G$, $m(E) = m(gE) \forall g \in G$, and this will suffice to show $|det A_x|^{-1} dx$ is a left Haar measure.

$$m(gE) = \int_{gE} |\det A_x|^{-1} dx$$

now let $x = gy = A_g(y) + b_g \implies dx = |\det A_g| dy$

$$\implies \int_E |\det A_{gy}|^{-1} |\det A_g| dy$$

Note that A_{gy} is a linear transformation over $\mathbb{R}^n$, and thus $A_{gy} = A_g A_y$

$$\begin{aligned} &= \int_E |\det A_g \det A_y|^{-1} |\det A_g| dy \\ &= \int_E |\det A_y|^{-1} dy = m(E) \end{aligned}$$

The same argument holds for the right translation of E. □

Problem 8 (Folland 11.4)**.** *The following are special cases of Exercise 3.*

(a) If G is the multiplicative group of nonzero complex numbers $z = x+iy$, $(x^2+y^2)^{-1}\, dx\, dy$ is a Haar measure.

(b) *If G is the group of invertible $n\times n$ real matrices, $|\det A|^{-n}\,dA$ is a left and right Haar measure, where $dA =$ Lebesgue measure on $\mathbb{R}^{n\times n}$. (To see that the determinant of the map $X \mapsto AX$ is $|\det A|^n$, observe that if X is the matrix with columns $X^1,\dots,X^n$, then AX is the matrix with columns $AX^1,\dots,AX^n$.)*

(c) *If G is the group of 3×3 matrices of the form*

$$\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}$$

where $(x,y,z\in\mathbb{R})$, then $dx\,dy\,dz$ is a left and right Haar measure.

(d) *If G is the group of 2×2 matrices of the form*

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$$

where $(x>0, y\in\mathbb{R})$, then $x^{-2}\,dx\,dy$ is a left Haar measure and $x^{-1}\,dx\,dy$ is a right Haar measure.

Proof. (a) Here, we have that $m(E)=\int_E \frac{1}{x^2+y^2}dxdy$, so define the translation $T(E)=wE$, where $E\subset\mathbb{C}$. We see that if $w=u+iv$ and $z=x+iy$, then $wz=ux-vy+i(uy+vx)$, so set $\tilde{x}=ux-vy$ and $\tilde{y}=uy+vx$. Moreover, the Jacobian of this transformation from $(x,y)\to(\tilde{x},\tilde{y})$ is $\frac{\partial\tilde{x}}{\partial x}\frac{\partial\tilde{y}}{\partial y}-\frac{\partial\tilde{x}}{\partial y}\frac{\partial\tilde{y}}{\partial x}=u^2+v^2 \implies d\tilde{x}d\tilde{y}=(u^2+v^2)dxdy$. Thus, we have

$$m(T(E))=\int_{wE}\frac{1}{x^2+y^2}dxdy=\int_E\frac{1}{\tilde{x}^2+\tilde{y}^2}d\tilde{x}d\tilde{y}$$

after where $\tilde{x},\tilde{y}$ are defined as above. Note that we can write $\tilde{x}^2+\tilde{y}^2=(xu-yv)^2+(xv+yu)^2=(u^2+v^2)(x^2+y^2)$. Finally, substituting $d\tilde{x}d\tilde{y}=(u^2+v^2)dxdy$ into the integral, we have

$$\int_E\frac{1}{\tilde{x}^2+\tilde{y}^2}d\tilde{x}d\tilde{y}=\int_E\frac{1}{(u^2+v^2)(x^2+y^2)}(u^2+v^2)dxdy=\int_E\frac{1}{x^2+y^2}dxdy$$

(b) Let $X\subset GL(n,\mathbb{R})$ be a Borel subset of $GL(n,\mathbb{R})$, and we desire to show that $m(AX)=\int_{AX}\frac{dA}{|\det A|^n}$. Now we can see that for $T: GL(n,\mathbb{R})\to GL(n,\mathbb{R})$ defined as the left translation $T(X)=AX$, $T(X_1\bigoplus X_2\bigoplus\cdots\bigoplus X_n)=AX_1\bigoplus AX_2\bigoplus\cdots\bigoplus AX_n$, and thus we see that the Jacobian of T is $\det(A)^n$ by applying the determinant through the tensor product. Now we preform a change of variables to obtain that

$$\int_{T(X)}\frac{dx}{|\det(x)|^n}=\int_X\frac{1}{|\det Au|^n}(\det A)^n du=\int_E \det(u)^{-n}du=m(E)$$

(c) Here, note that for left multiplication, fix a matrix $M\in G$ such that $M=\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1\end{pmatrix}$ and for $g\in g$,

$$Mg=\begin{pmatrix} 1 & a+x & az+b+y \\ 0 & 1 & c+z \\ 0 & 0 & 1\end{pmatrix}$$

and

$$gM = \begin{pmatrix} 1 & a+x & b+cx+y \\ 0 & 1 & c+z \\ 0 & 0 & 1 \end{pmatrix}$$

in both cases, $dxdydz$ is invariant under these translations, since $\det Mg = \det gM = \det g = 1$, and hence the volume is preserved under both left and right translation. So indeed, $dxdydz$ is a Haar Measure.

(d) Here, we see that for $M = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$, and for $g \in G$

$$Mg = \begin{pmatrix} ax & ay+1 \\ 0 & 1 \end{pmatrix}$$

and

$$gM = \begin{pmatrix} ax & by+1 \\ 0 & 1 \end{pmatrix}$$

So we see that left multiplication scales the matrix g by a factor of a^2, and right multiplication scales by a factor of a. The Haar Measures given correctly account for this scaling factor, and are left and right invariant, respectively. □

Problem 9. *Let $\{z_n\}_{n=1}^{\infty}$ be a sequence of complex numbers, and let $s_n = \sum_{k=1}^{n} z_k$ be partial sums. Let $\sigma_n = \frac{1}{n}\sum_{k=1}^{n} s_k$.*

(a) Show that if $s_n \to s$, then $\sigma_n \to s$.

(b) Give an example of $\{z_n\}$ for which σ_n converges but s_n does not.

(c) Assume that $\sigma_n \to \sigma$ and $nz_n \to 0$, then $s_n \to \sigma$. [Hint: show directly that $s_n - \sigma_n \to 0$.]

Proof. (a) Given $s_n \to s$, $\exists \epsilon > 0$ s.t. $|s_n - s| < \epsilon n \geq N$, for some $N \in \mathbb{N}$. Note that

$$\begin{aligned} |s_n - \sigma_n| &= |\sum_{k=1}^{n} z_n - \frac{1}{n}\sum_{k=1}^{n} s_n| \\ &= |s_n - \frac{1}{n}\sum_{k=1}^{N} s_k - \frac{1}{n}\sum_{k=N+1}^{n} s_k| \end{aligned}$$

now as $n \to \infty$, we have that $s \to s_n$, $\frac{1}{n}\sum_{k=1}^{N} s_k \to 0$ as $\sum_{k=1}^{N} s_k$ is some finite sum not depending on n, and $\frac{1}{n}\sum_{k=N+1}^{n} s_k \to s_n \to s$, and this $|s_n - \sigma_n| \to |s - s_n| < \epsilon \implies \sigma_n \to s_n \to s$ which is what we desired to show.

(b) Here, we can choose $z_n = \frac{1}{n}$. It is clear that

$$\lim_{n\to\infty} s_n = \sum_{k=1}^{\infty} \frac{1}{n}$$

diverges as the harmonic series. However note that

$$\ln(n) < s_n < \ln(n+1)$$
$$\implies \frac{1}{n}\sum_{k=1}^{n}\ln(n) < \sigma_n < \frac{1}{n}\sum_{k=1}^{n}[\ln(k)+1]$$

Now notice that $\frac{1}{n}\sum_{k=1}^{n}\ln(n) = \frac{1}{n}\ln(\prod_{k=1}^{n}k) = \frac{1}{n}\ln(n!) \approx \frac{1}{n}\ln(\sqrt{2\pi n}(n/e)^n) = ln(n/e)+\frac{ln(\sqrt{2\pi n})}{n}$ by Stirling's formula. Thus, asymptotically, we have that $\ln(n)-1 < \sigma_n < \ln(n)+1$. Thus, as $n\to\infty$, $\ln(n)-1$ and $\ln(n)+1$ both diverge, but their difference remains constant. Thus, σ_n approaches a definite trend, and converges as a sum of Cesaro terms. In fact, $\lim_{n\to\infty}\sigma_n = \gamma$, where γ is the Euler Mascheroni constant, though this is much harder to prove.

(c) First note that we can write $\sigma_n = \frac{1}{n}(s_1+s_2+...+s_n) = \frac{1}{n}(z_1+(z_1+z_2)+...) = \frac{1}{n}(nz_1+(n-1)z_2+(n-2)z_3+...+2z_{n-1}+z_n)$. Thus, we have that

$$s_n - \sigma_n = \sum_{k=1}^{n} z_k - \frac{1}{n}(nz_1+(n-1)z_2+(n-2)z_3+...+2z_{n-1}+z_n)$$
$$= \frac{1}{n}z_2 + \frac{2}{n}z_3 + ... + \frac{n-2}{n}z_{n-1} + \frac{n-1}{n}z_n$$

Since we assume that $nz_n \to 0$, then each of the terms converges to zero as $n\to\infty$. Thus, we have that $s_n - \sigma_n \to 0 \implies s_n \to \sigma_n \to \sigma$.

□

Problem 10. *This question is about convolutions on $\mathbb{R}^n$:*

$$f * g(x) = \int f(y)g(x-y)dy.$$

(a) Show that if $f \in L^1$ and $g \in L^p$, $1 \le p \le \infty$, then $f*g \in L^p$ and

$$\|f*g\|_{L^p} \le \|f\|_{L^1}\|g\|_{L^p}.$$

(b) Show that for $f,g,h \in L^1$, $f*g = g*f$ and $(f*g)*h = f*(g*h)$.

Proof. (a) Here, we have that $||f*g||_{L^p} = (\int_{\mathbb{R}^n}|\int_{\mathbb{R}^n}f(y)g(x-y)dy|^p dx)^{1/p}$. Now by Minkowski's integral inequality, we can switch the powers and order of integration to obtain:

$$(\int_{\mathbb{R}^n}|\int_{\mathbb{R}^n}f(y)g(x-y)dy|^p dx)^{1/p} \le \int_{\mathbb{R}^n}(\int_{\mathbb{R}^n}|f(x)g(x-y)|^p dx)^{1/p}dy$$

Now by the translation invariance of the L_p norm, we have that

$$\int_{\mathbb{R}^n}(\int_{\mathbb{R}^n}|f(x)g(x-y)|^p dx)^{1/p}dy \le \int_{\mathbb{R}^n}|f(y)|||g||_{L^p}dy = ||g||_{L^p}\int_{\mathbb{R}^n}|f(y)|dy = ||f||_{L^1}||g||_{L^p}$$

and thus $||f*g||_{L^p} \le ||f||_{L^1}||g||_{L^p}$

(b) First note that $f * g = \int_{\mathbb{R}^n} f(y)g(x-y)dy$ and after the change of variables $z = x - y \implies dy = dz$ we have $\int_{\mathbb{R}^n} f(x-z)g(z)dy$ and this is precisely the definition of $g * f$. For associativity, note that

$$((f*g)*h)(x) = \int_{\mathbb{R}^n}(\int_{\mathbb{R}^n} f(z)g(y-z)dz)h(x-y)dy = \int_{\mathbb{R}^n} f(z)(\int_{\mathbb{R}^n} g(y-z)h(x-y)dy)dz$$

by Fubini's Theorem. Now we can make the substitution $v = x - y$ to obtain

$$\int_{\mathbb{R}^n} f(z)(\int_{\mathbb{R}^n} g(x-z-v)h(v)dv)dz = \int_{\mathbb{R}^n} f(z)(g*h)(x-z)dz = (f*(g*h))(x)$$

and thus $((f*g)*h)(x) = (f*(g*h))(x)$

□

Problem 11. *Let $f \in C^1(\mathbb{T})$. Show that $\lim_{|n|\to\infty} n|\hat{f}(n)| = 0$.*

Proof. Here, we have that $\hat{f}(n) = \int_{\mathbb{T}} f(x)e^{-2\pi inx}dx$. Since we are told that $f \in C^1$, we can employ integration by parts and differentiate f: set $u = f \implies u' = f'$ and $v' = e^{-2\pi inx} \implies v = -\frac{e^{-2\pi inx}}{2\pi in}$. This yields

$$\hat{f} = -\frac{f(x)e^{-2\pi inx}}{2\pi in}|_{\mathbb{T}} + \frac{1}{2\pi in}\int_{\mathbb{T}} f'(x)e^{2\pi inx}dx = \frac{1}{2\pi in}\int_{\mathbb{T}} f'(x)e^{2\pi inx}dx$$

since f is 1-periodic so its boundary points cancel out in the first term. Now we have that $\lim_{|n|\to\infty} n|\hat{f}(n)| = \lim_{|n|\to\infty} \frac{1}{2\pi i}\int_{\mathbb{T}} f'(x)e^{-2\pi inx}dx$. Now since $f'(x), e^{-2\pi inx} \in L^1(\mathbb{T})$, we can employ Lebesgue Dominated Convergence to pass the limit through the integral, and hence we obtain

$$\lim_{|n|\to\infty} \frac{1}{2\pi i}\int_{\mathbb{T}} f'(x)e^{-2\pi inx}dx = \frac{1}{2\pi i}\int_{\mathbb{T}} \lim_{|n|\to\infty} f'(x)e^{-2\pi inx}dx = 0$$

as $\lim_{|n|\to\infty} e^{-2\pi inx} = 0$. Therefore, $\lim_{|n|\to\infty} n|\hat{f}(n)| = 0$ which is what we desired to show.

□

Problem 12. *Let $f(x) = \min\{x, 1-x\}$ (extended to be 1-periodic from $[0,1]$).*

1. *Compute $\hat{f}$.*
2. *Use this to compute $\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}$ and $\sum_{n=1}^{\infty} \frac{1}{(2n-1)^4}$.*

Proof. (a) Here, we consider the Fourier transform as two integrals that represent our given $f(x)$ on the intervals $[0,1/2]$ and $[1/2,1]$ respectively.

$$\implies \hat{f} = \int_0^{1/2} xe^{-2\pi inx}dx + \int_{1/2}^1 (1-x)e^{-2\pi inx}dx$$

and we see that after a using integration by parts and summing both integrals, we have that

$$\hat{f} = \frac{\sin^2(\frac{\pi n}{2})\cos(\pi n)}{\pi^2 n^2} + i\frac{sin(\pi n)(\cos(\pi n) - 1)}{2\pi^2 n^2}$$

(b) Here, we have that $\Re(\hat{f}(2n+1)) = \frac{1}{\pi^2(2n+1)^2}$. Now using the fact that $\int_0^1 f(x)dx = \sum_{n=-\infty}^{\infty} \hat{f}(n)$, we first evaluate that $\int_0^1 f(x)dx = \int_0^{1/2} xdx + \int_{1/2}^1 (1-x)dx = 1/4$. So now

$$1/4 = \frac{1}{\pi^2}\sum_{n=-\infty}^{\infty}\frac{1}{(2n+1)^2} \implies \sum_{n=1}^{\infty}\frac{1}{(2n+1)^2} = \frac{\pi^2}{8} - 1$$

A similar calculation yields that

$$\sum_{n=1}^{\infty}\frac{1}{(2n-1)^4} = \frac{\pi^4}{96} - 1$$

□

Problem 13. *Let $L \in GL_n(\mathbb{R})$, and $f \in L^1(\mathbb{R}^n)$.*

(a) If $g(x) = f(Lx)$, find the relation between $\hat{f}$ and $\hat{g}$.

(b) Specialize to the case of $L \in SO_n(\mathbb{R})$ to show that if f is radial, so is $\hat{f}$.

Proof. (a)

$$\hat{g}(\xi) = \int g(x)e^{2\pi i x\cdot\xi}dx = \int f(Lx)e^{2\pi i x\cdot\xi}dx$$

Now we perform the change of variables $Lx = y \implies |\det L|dx = dy \implies dx = \frac{1}{|\det L|}dy$ to obtain

$$\begin{aligned} &\int f(y)e^{2\pi i(L^{-1}y)\cdot\xi}\frac{1}{|\det L|}dy \\ &= \frac{1}{|\det L|}\int f(y)e^{2\pi i y\cdot(L^{-1})^T\xi}dy \\ &= \frac{1}{|\det L|}\hat{f}((L^{-1})^T\xi) \\ \implies &\boxed{\hat{g}(\xi) = \frac{1}{|\det L|}\hat{f}((L^{-1})^T\xi)} \end{aligned}$$

(b) f is radial, and hence $f(x)$ depends on the norm of x alone for any $x \in \mathbb{R}^n$. Now since orthogonal matrices preserve distance, that is, for any $A \in O_n(\mathbb{R})$, $\|Ax\| = \|x\|$. Therefore, for f radial, $f(Ax) = f(x)$. Now assume $L \in SO_n(\mathbb{R})$, and observe

$$\hat{f}(L\xi) = \int f(x)e^{2\pi i x\cdot(L\xi)}dx = \int f(x)e^{2\pi i(Lx)\cdot\xi}dx$$

Now perform the change of variables $Lx = y \implies x = L^{-1}y = L^Ty \implies |\det L|dx = dy \implies dx = dy$ since $L \in SO_n(\mathbb{R})$ and hence $\det L = 1$.

$$\implies \hat{f}(L\xi) = \int f(L^Ty)e^{2\pi i y\cdot\xi}dy = \int f(y)e^{2\pi i y\cdot\xi}dy = \hat{f}(\xi)$$

since L^T is also orthogonal so $f(L^Ty)) = f(y)$ and thus $\hat{f}(L\xi) = \hat{f}(\xi)$, and $L \in SO_n(\mathbb{R}) \subset O_n(\mathbb{R})$ so $\hat{f}$ is indeed radial.

□

Problem 14. *Let $f : \mathbb{R} \to \mathbb{C}$ be a continuously differentiable function with $|f| + |f'| \leq C(1+|x|)^{-\alpha}$ for some $C, \alpha > 0$.*

(a) Find a formula for $\hat{f}'$ in terms of $\hat{f}$.

(b) Show that

$$\int_{\mathbb{R}} |f|^2 \leq 2\sqrt{\int_{\mathbb{R}} x^2|f(x)|^2 dx \int_{\mathbb{R}} |f'(x)|^2 dx}.$$

[Hint: integrate by parts to get $x(|f|^2)'$.]

(c) Show that

$$\left(\int_{\mathbb{R}} |f|^2\right)^2 \frac{1}{16\pi^2} \leq \int_{\mathbb{R}} x^2|f(x)|^2 dx \int_{\mathbb{R}} \xi^2|\hat{f}(\xi)|^2 d\xi.$$

In other words, if $\int_{\mathbb{R}} |f|^2$ is fixed, there is a limit to how much both f and $\hat{f}$ can be localized near the origin.

(d) Find all the f for which

$$\left(\int_{\mathbb{R}} |f|^2\right)^2 \frac{1}{16\pi^2} = \int_{\mathbb{R}} x^2|f(x)|^2 dx \int_{\mathbb{R}} \xi^2|\hat{f}(\xi)|^2 d\xi.$$

Proof. (a) Note that $\hat{f}'(\xi) = \int_{-\infty}^{\infty} f'(x)e^{-2\pi i x\xi}dx$. Here, we perform integration by parts, setting $u = e^{2\pi i x\xi}$ and $v' = f'(x)$, so we have

$$\begin{aligned} \hat{f}'(\xi) &= f(x)e^{-2\pi i x\xi}|_{-\infty}^{\infty} + \int 2\pi i\xi f(x)e^{-2\pi i x\xi}dx \\ &= 2\pi i\xi \int f(x)e^{-2\pi i x\xi}dx \\ &= 2\pi i\xi \hat{f}(\xi) \end{aligned}$$

since $|f| + |f'| \leq C(1+|x|)^{-1-\alpha}$, and hence f vanishes at ∞ and $f(x)e^{-2\pi i x\xi}|_{-\infty}^{\infty} = 0$. Therefore, $\hat{f}'(\xi) = 2\pi i\xi\hat{f}(\xi)$.

(b) Using integration by parts, we set $u = |f|^2$ and $v' = 1$, and we have that $u' = (|f|^2)'$ and $v = x$. Thus,

$$\int_{\mathbb{R}} |f|^2 dx = (x|f|^2)|_{-\infty}^{\infty} - \int_{\mathbb{R}} x(|f|^2)'dx$$

Now by Cauchy-Schwarz, we have that

$$(\int_{\mathbb{R}} x(|f|^2)'dx)^2 \leq \int_{\mathbb{R}} x^2 dx \int ((|f|^2)')^2 dx \leq 4\int_{\mathbb{R}} |x|^2 f(x)^2 \int_{\mathbb{R}} (f'(x))^2 dx$$

Taking square roots on both sizes, we obtain

$$\int_{\mathbb{R}} |f|^2 \leq 2\sqrt{\int_{\mathbb{R}} x^2|f(x)|^2 dx \int_{\mathbb{R}} |f'(x)|^2 dx}$$

as desired.

(c) Using the previous inequality and squaring both sides, we obtain that

$$(\int_{\mathbb{R}} |f|^2)^2 \leq 4 \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} |f'(x)|^2 dx$$

Now note that due to the decay condition given by $|f| + |f'| \leq C(1+|x|)^{-\alpha-1}$, as $|x| \to \infty$, $C(1+|x|)^{-\alpha-1}0$, and hence $|f| + |f'| \to 0 \implies |f|, |f'| \to 0$. Given this decay condition, we have that $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and we can employ Plancherel's Theorem:

$$\begin{aligned}
(\int_{\mathbb{R}} |f|^2)^2 &\leq 4 \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} |f'(x)|^2 dx \\
&= 4 \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} |\hat{f}'(\xi)|^2 d\xi \\
&= 4 \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} |2\pi i \xi \hat{f}(\xi)|^2 d\xi \\
&= 16\pi^2 \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} |\xi|^2 |\hat{f}(\xi)|^2 d\xi \\
\implies (\int_{\mathbb{R}} |f|^2)^2 \frac{1}{16\pi^2} &\leq \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} |\xi|^2 |\hat{f}(\xi)|^2 d\xi
\end{aligned}$$

(d) Note that the inequality obtained from part (b) is Holder's inequality for L^2 functions. In order for this inequality to be an equality, we require that the two functions $g = x$ and $h = (|f|^2)'$ are linearly independent w.r.t the L^2 norm, that is $g^2 = \lambda h^2 \implies x^2 = \lambda((|f|^2)')^2$. We see that setting $f(x) = \rho x$, where $\rho \in \mathbb{R}$ satisfies this condition for linear independence:

$$x^2 = \lambda((|f|^2)')^2 \implies x^2 = \lambda((|\rho x|^2)')^2 = (2\rho x)^2 = 4\rho^2 x^2 = \lambda x^2$$

after setting $\lambda = 4\rho^2$. Thus, the family of functions satisfying this condition are given by

$$\mathcal{F} = \{\rho x : \rho \in R\}$$

$\square$

Problem 15. *Assume that $f \in L^1(\mathbb{R})$ is a function with Fourier Transform supported on B_R.*

(a) Show that $f \in L^\infty(\mathbb{R}^n)$ and

$$\|f\|_{L^\infty(\mathbb{R}^n)} \leq CR^n \|f\|_{L^1(\mathbb{R}^n)},$$

and that this is the only possible dependence on R in such an estimate (i.e. there is a family of functions f_R with $\hat{f}_R$ supported on B_R for which $R^n \|f_R\|_{L^1} \leq c\|f_R\|_{L^\infty}$).

(b) Show that $\hat{f} \in C^\infty(\mathbb{R}^n)$ and

$$\sup_{\mathbb{R}^n} |D^k f| \leq (CR)^n \|f\|_{L^1(\mathbb{R}^n)}.$$

[More precisely, show that f coincides with a smooth function almost everywhere.]

(c) *Conclude that f is real analytic (its Taylor series around any point has positive radius of convergence and converges to f).*

Proof. (a) Since $f \in L^1(\mathbb{R})$, we can use the Fourier Inversion formula, which states that

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi$$

and derive the following inequality since $\hat{f}$ is supported on B_R:

$$\begin{aligned} f(x) &= \int_{B_R} \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi \\ \implies |f(x)| &\leq \int_{B_R} |\hat{f}(\xi)| \cdot |e^{2\pi i \xi \cdot x}| d\xi = \int_{B_R} |\hat{f}(\xi)| d\xi \\ \implies \|f\|_{L^\infty(\mathbb{R}^n)} &= \sup_{x \in \mathbb{R}^n} |f(x)| \leq \sup_{x \in \mathbb{R}^n} \int_{B_R} |\hat{f}(\xi)| d\xi \\ \implies \|f\|_{L^\infty(\mathbb{R}^n)} &\leq CR^n \|\hat{f}\|_{L^\infty(\mathbb{R}^n)} \\ \implies \|f\|_{L^\infty(\mathbb{R}^n)} &\leq CR^n \|f\|_{L^1(\mathbb{R}^n)} \end{aligned}$$

(b) Since the Fourier Transform of f is compactly supported, by the Paley-Weiner Theorem, f can be extended to an entire function on $\mathbb{C}^n$, which suffices to show that $f \in C^\infty(\mathbb{R}^n)$. Moreover, we have the k-th derivative of f is given by $D^k(f) = (2\pi i \xi)^k \hat{f}(\xi)$ and thus we have that

$$\sup_{x \in \mathbb{R}^n} |D^k f(x)| \leq \int_{B_R} |(2\pi i \xi)^k \hat{f}^k| d\xi$$

but $\hat{f}$ is bounded and hence $(2\pi i \xi)^k$ grows as $R^|k|$, and the integral is bounded by a constant C times $R^{n+|k|}$, as so

$$\int_{B_R} |(2\pi i \xi)^k \hat{f}^k| d\xi \leq (CR)^{|k|} \|\hat{f}\|_{L^1(\mathbb{R}^n)}$$

for the appropriate choice of C.

(c) First, $f \in C^\infty(\mathbb{R}^n)$, and the Taylor series expansion of f is given by $f(x+h) \sum_k \frac{D^k f(x)}{k!} h^k$, and using the bound for the $k-th$ derivative of f, we have that the remainder term of this taylor series expansion of f converges to 0, and hence f has a convergent power series expansion at every $x \in \mathbb{R}^n$, and so f is real-analytic. □

Problem 16. *Show that if $f \in L^1(\mathbb{R}^n)$ and both f and $\hat{f}$ have compact support, then $f = 0$ almost everywhere.*

Proof. Here we can appeal to 15(c). Since $\hat{f}$ has compact support, f is real-analytic and can be extended to an analytic function on all of $\mathbb{C}^n$ as a consequence of Paley-Weiner theorem. Now f also has compact support, and thus f is identically zero on $U = \mathbb{R}^n \setminus \text{supp} f$. By the identity theorem, since U a non-empty open set in $\mathbb{C}^n$ and hence contains (infinitely many) accumulation points and $f|_U = 0$, then f must be zero on its entire domain of analyticity, which is all of $\mathbb{R}^n$. Therefore, $f \equiv 0$ a.e. □

Problem 17. *Consider the partial differential equation*

$$-\Delta u + u = f \text{ on } \mathbb{R}^n,$$

where $\Delta u = div(\nabla u) = Tr(D^2 u)$.

(a) For $f \in \mathcal{S}$ *(the Schwartz space), show that there exists a unique* $u \in \mathcal{S}$ *which solves this equation.*

(b) Show that, for $f \in \mathcal{S}$ *and* u *the solution you found,*

$$\sum_{|\alpha|\leq 2} \|\partial^\alpha u\|_{L^2(\mathbb{R}^n)} \leq C\|f\|_{L^2(\mathbb{R}^n)}$$

for a $C > 0$ *depending only on* n.

(c) (Optional) Is this still true with higher-order derivatives on the left?

Proof. (a) Here, we can apply the Fourier Transform to both sides of the equation. The Fourier Transform of the Laplacian Δu is $-|\xi|^2\hat{u}(\xi)$, and thus

$$-\Delta u + u = f \implies |\xi|^2\hat{u}(\xi) + \hat{u}(\xi) = \hat{f}(\xi)$$

now solving for $\hat{u}$, we have that

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{|\xi|^2 + 1}$$

and now we take the inverse Fourier transform of u. Since $f \in \mathcal{S}$, $\frac{\hat{f}(\xi)}{|\xi|^2+1} \in \mathcal{S}$ since the denominator moderates any growth of $\hat{f}$. After taking the inverse Fourier Transform, we obtain that $u \in \mathcal{S}$ since the denominator $(|\xi|^2 + 1)^{-1}$ ensures that u decays rapidly enough, as any derivatives $D^\alpha u$ will decay faster than u. Moreover, $u \in C^\infty(\mathbb{R}^n)$ since the denominator of $D^\alpha u$ is never zero. For uniqueness: say that $v = u_1 - u_2$ are two solutions of the given differential equation, then letting $f = 0$ yields

$$-\Delta v + v = 0 \implies (|\xi|^2 + 1)\hat{v}(\xi) = 0 \implies v = 0 \implies u_1 = u_2$$

(b) Given that $D^\alpha(u) = (i\xi)^\alpha\hat{u}(\xi)$, we have that

$$\|D^\alpha(u)\|_{L^2} = \|(i\xi)^\alpha\hat{u}(\xi)\|_{L^2} = (\int_{\mathbb{R}^n} ((i\xi)^\alpha\hat{u}(\xi))^2 d\xi)^{1/2})$$

now replacing $\hat{u} = \frac{\hat{f}}{|\xi|^2+1}$, we have that

$$(\int_{\mathbb{R}^n} ((i\xi)^\alpha\hat{u}(\xi))^2 d\xi)^{1/2}) = \int_{\mathbb{R}^n} ((i\xi)^\alpha(\frac{\hat{f}}{|\xi|^2 + 1})(\xi))^2 d\xi)^{1/2}$$

Now since $|\alpha| \leq 2$, $(i\xi)^\alpha$ is at most of order $|\xi|^4$, we have that

$$\|D^\alpha(u)\|_{L^2} \leq (\int_{\mathbb{R}^n} \frac{|\xi|^4|\hat{f}|^2}{(|\xi|^2+1)^2} d\xi)^{1/2}$$

and since $|\xi+1|^2 \leq |\xi|^4$, we have that

$$\|D^\alpha(u)\|_{L^2} \leq (\int_{\mathbb{R}^n} |\hat{f}|^2 d\xi)^{1/2} = \|f\|_{L^2}$$

Passing through sums, we have

$$\sum_{|\alpha|\leq 2} \|D^\alpha(u)\|_{L^2} \leq (\sum_{|\alpha|\leq 2} 1)\|f\|_{L^2}$$

setting $C = (\sum_{|\alpha|\leq 2} 1)$, which depends only on N, we obtain the desired inequality, namely:

$$\sum_{|\alpha|\leq 2} \|D^\alpha(u)\|_{L^2} \leq C\|f\|_{L^2}$$

(c) No, since $(i\xi)^\alpha$ would be of order $|\xi|^k$, where $k > 4$, which cannot guarantee the decay condition in the integral used in part (b).

□

Problem 18. *Consider the partial differential equation*

$$-\Delta u = f \text{ on } \mathbb{R}^n.$$

(a) Assume $f \in C_c^\infty(\mathbb{R}^n)$. Show that if $n > 2$, there exists a solution $u \in C^\infty(\mathbb{R}^n)$ for which u and all of its derivatives are bounded.

(b) Show that for the solution you found, $|u(x)| \leq \frac{C}{|x|^{n-3}}$ (where the constant depends on f). [Hint: take derivatives on the Fourier side.]

(c) (Optional) Show that for the solution you found, in fact $|u(x)| \leq \frac{C}{|x|^{n-2}}$.

(d) (Optional) For every $k > n-2$, find explicit conditions on f such that the solution u has $|u(x)| \leq \frac{C}{|x|^k}$.

Proof. (a) $-\Delta u = f \implies |\xi|^2\hat{u}(\xi) = \hat{f}(\xi) \implies \hat{u}(\xi) = \frac{\hat{f}(\xi)}{|\xi|^2}$. Now since $f \in C_c^\infty(\mathbb{R}^n)$, $\hat{f}(\xi)$ decays rapidly and vanishes at ∞. Since f has compact support and $\hat{f}$ is smooth, $\hat{u}$ vanishes at least as fast as $\frac{1}{|\xi|^2}$, and hence $\hat{u}$ is a tempered distribution. Now u is given by

$$u(x) = \mathcal{F}^{-1}(\frac{\hat{f}(\xi)}{|\xi|^2})$$

and since $\mathcal{F}^{-1}$ of a smooth rapidly decreasing function is itself a smooth function, and so $u(x) \in C^\infty$. The derivatives of u are all bounded since $\hat{u}$ is rapidly decreasing.

(b) It is clear that away from 0, $\hat{u}(\xi) = \frac{\hat{f}(\xi)}{|\xi|^2}$ decreases rapidly. Note that $\hat{f}$ is smooth, and hence can be expressed as Taylor series around $\xi = 0$. We have that $\hat{f}(\xi) \approx \hat{f}(0) + O(|\xi|)$, and since f has integral 0 as a compactly supported function with at least one non-vanishing derivative, $\hat{f}(0) = \int f(x)dx = 0$, and hence $\hat{f}(0) = O(|\xi|)$ as $\xi \to 0$. Therefore, as $\xi \to 0$, $\hat{u}(\xi)$ behaves like $O(|\xi|^{-1})$. Now the inverse Fourier transform in $\mathbb{R}^n$ scales like $\frac{1}{|x|^{n-2}}$ when the Fourier transform is $O(|\xi|^{-1})$. Due to the presence of a $|\xi|^2$ in the denominator of $\hat{u}$, this decreases the decay rate of $u(x)$ by a factor of $\frac{1}{|x|}$. Therefore, $u(x)$ behaves like $\frac{1}{|x|^{n-3}}$ for large n, and so

$$u(x) \leq \frac{C}{|x|^{n-3}}$$

for some $C > 0$

□

Problem 19. *For $u : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$, consider the heat equation*

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{on } \mathbb{R}^n \times (0, \infty) \\ u(x, 0) = u_0(x) \end{cases}$$

(a) For $u_0 \in \mathcal{S}$, find a candidate solution u using the Fourier transform in the x variables only, and give a formula for u in the form $u(x,t) = u_0 * K_t$, where $K_t(x) : \mathbb{R}^n \to \mathbb{R}$ is explicit.

Hint: we saw this in class; solve ODEs on the Fourier side.

(b) Show that the u given by this formula is in $\mathcal{S}$ for every t, is continuously differentiable in t for $t > 0$, and actually solves $\partial_t u = \Delta u$ for $t > 0$.

(c) Show that the u is continuous up to $t = 0$ and $u(x, 0) = u_0(x)$

Hint: recall our theorems about approximate identities.

Proof. (a) Here, we apply the Fourier Transform to the heat equation to obtain:

$$\partial_t u - \Delta u = 0 \implies \partial_t \hat{u}(\xi, t) - 4\pi^2|\xi|^2\hat{u}(\xi, t) = 0$$

Now we see that this is the ODE

$$\frac{d\hat{u}}{dt} + 4\pi^2|\xi|^2\hat{u} = 0$$

which has the solution

$$\hat{u}(\xi, t) = \hat{u}_0(\xi)e^{-4\pi^2|\xi|^2 t}$$

Now in the Fourier space, the expression $e^{-4\pi^2|\xi|^2 t}$ corresponds to the heat kernel given by

$$K_t(x) = \frac{1}{(4\pi t)^{n/2}}e^{-\frac{|x|^2}{4t}}$$

which has the Fourier transform

$$\hat{K}_t(\xi) = e^{-4\pi^2|\xi|^2 t}$$

Now note that

$$u(x,t) = \mathcal{F}^{-1}[\hat{u}(\xi,t)](x)$$

and the Fourier transform of a convolution is the product of Fourier transforms, and so

$$u(x,t) = u_0 \circledast K_t(x) = \int_{\mathbb{R}^n} u_0(y) K_t(x-y)dy$$

as required.

(b) First, note that $u_0 \in \mathcal{S}$, and $K_t(x) \in C^\infty(\mathbb{R}^n)$ and its derivatives decay exponentially fast. Hence,

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial}{\partial t}\int_{\mathbb{R}^n} u_0(y)K_t(x-y)dx = \int_{RR^n} \frac{\partial}{\partial t}[u_0(y)K_t(x-y)]dx$$

differentiation w.r.t. t of the heat kernel is smooth for the reasons given, so indeed u is continuously differentiable under t. Lastly, note that

$$\partial_t u(x,t) = \int_{\mathbb{R}^n} u_0(y)\frac{\partial}{\partial t}K_t(x-y)dy$$

and

$$\Delta u(x,t) = \int_{\mathbb{R}^n} u_0(y)\Delta K_t(x-y)dy$$

so it will suffice to show that $\Delta K_t(x-y) = \frac{\partial}{\partial t}K_t(x-y)$. We have that

$$\frac{\partial}{\partial t}K_t(x-y) = \frac{\partial}{\partial t}[\frac{1}{(4\pi t)^{n/2}}e^{-\frac{|x|^2}{4t}}] = (-\frac{n}{2\pi(4\pi t)^{n/2}} + \frac{|x|^2}{4t^2(4\pi t)^{n/2}})(e^{-\frac{|x|^2}{4t}})$$

and

$$\Delta K_t(x-y) = \mathcal{F}^{-1}(-4\pi^2|\xi|^2 e^{-4\pi^2|xi|^2 t}) = \frac{\partial}{\partial t}K_t(x)$$

by the properties of Fourier inversion and convolutions and since $\Delta e^{-\frac{|x|^2}{4t}}$ in the Fourier domain corresponds to multiplication by $-|\xi|^2$. Thus $\Delta K_t(x-y) = \frac{\partial}{\partial t}K_t(x-y)$, and so $\Delta u(x,t) = \frac{\partial}{\partial t}u(x,t)$.

(c) We want to show that

$$\lim_{t\to 0^+} u(x,t) = \lim_{t\to 0^+}\int_{RR^n} u_0(y)K_t(x-y)dy = u_0(t)$$

Given the fact that $K_t(x) \in C^\infty(\mathbb{R}^n)$ and decays rapidly except around of neigborhood of $y = x$, by the dominated convergence theorem, we have that

$$\lim_{t\to 0^+}\int_{RR^n} u_0(y)K_t(x-y)dy = \int_{RR^n}\lim_{t\to 0^+} u_0(y)K_t(x-y)dy$$

now the heat kernel $K_t(x-y)$ behaves like a mollifier, meaning that it converges to the Dirac delta function $\delta(x-y)$ in the sense of distributions. Therefore,

$$\int_{RR^n} \lim_{t\to 0^+} u_0(y)K_t(x-y)dy = \int_{RR^n} u_0(y)\delta(x-y)dy = u_0(x)$$

and thus

$$\lim_{t\to 0^+} u(x,t) = u_0(t)$$

as required.

□

Problem 20. *We have seen that there are two basic estimates on the Fourier transform (of, say, a function $f \in \mathcal{S}$):*

$$\|\hat{f}\|_\infty \leq \|f\|_1$$
$$\|\hat{f}\|_2 \leq \|f\|_2.$$

It is possible to interpolate between these to get that

$$\|\hat{f}\|_q \leq \|f\|_p$$

where $1 \leq p \leq 2$ and $\frac{1}{p} + \frac{1}{q} = 1$ (e.g. Folland 8.30). Consider, now, whether an inequality like

$$\|\hat{f}\|_q \leq C_{p,q}\|f\|_p \tag{1}$$

is possible for other values of $p, q \in [1, \infty]$.

(a) Use the behavior of the Fourier transform under rescaling to show that if (0.1) is true, then $\frac{1}{p} + \frac{1}{q} = 1$.

(b) Let $\phi(x) = e^{-\pi|x|^2}$, and take for any $R > 0$ the sequence $f_n(x) = \phi(x + nR)e^{2\pi inRx}$. Show that $\|f_n\|_p = \|f\|_p = c_p$ do not depend on n or R.

(c) Show that, for each N,

$$2Nc_p = 2\sum_{n=1}^{N} \|f_n\|_p \geq \left\|\sum_{n=1}^{N} f_n\right\|_p \geq \frac{1}{2}N^{1/p}c_p$$

(d) If R is large enough. [Optional: you can take R independent of N.] Use the show that (0.1) is only possible if $q \geq 2p$, so the result in Folland 8.30 cannot be extended.

Proof. (a) Consider the scaling of a function f given by $f_\lambda(x) = f(\lambda x)$, and note that $\hat{f}(x) = (\frac{1}{\lambda^n})\hat{f}(\frac{\xi}{\lambda^n})$ for $f \in L^1(\mathbb{R}^n)$. In the L^p norm, we have that

$$\|f_\lambda\|_p^p = \int_{\mathbb{R}^n} |f(\lambda x)|^p dx = \frac{1}{\lambda^p}\|f\|_p^p$$

Now if we apply the given inequality, we have

$$(\int_{\mathbb{R}^n} |\frac{1}{\lambda^n}\hat{f}(\frac{\xi}{\lambda})|^q d\xi)^{1/q} \leq C_{p,q}(\frac{1}{\lambda^n}\int_{\mathbb{R}^n} |f(x)|^p dx)^{1/p}$$

now after making the change of variables $\xi' = \xi/\lambda$, we have that

$$\frac{1}{\lambda^{n/q}}(\int_{\mathbb{R}^n} |f(\xi')|^q d\xi')^{1/q} \leq C_{p,q}\lambda^{-n/p}\|f\|_p$$

and simplifying further, we have that

$$\lambda^{n(1/p-1/q)}\|\hat{f}\|_q \leq C_{p,q}\|f\|_p$$

For this inequality to hold, we must have that

$$n(\frac{1}{p}-\frac{1}{q}) = 0 \implies \frac{1}{p}-\frac{1}{q} = 0 \implies \frac{1}{p}+\frac{1}{q} = 1$$

where p and q are Hölder conjugates.

(b) Here, note that

$$\|f_n\|_p^p = \int_{\mathbb{R}} e^{-\pi p|x+nR|^2} dx$$

since this is a Gaussian integral, the translation by nR in the exponent does not effect the integral over $\mathbb{R}$, and hence

$$\|f_n\|_p^p = \int_{\mathbb{R}} e^{-\pi p|x|^2} dx$$

and this integral is evaluated as

$$\int_{\mathbb{R}} e^{-\pi p|x|^2} dx = (\frac{1}{\sqrt{p}})^n \int_{\mathbb{R}} e^{\pi x^2} dx$$

and hence

$$\|f_n\|_p = (\frac{1}{p^{n/2p}})(\frac{1}{\sqrt{\pi}})^{n/p} = c_p$$

and c_p does not depend on the choice of n or R.

(c) Here, denote $S_N(x) = \sum_{n=1}^N f_n(x)$. From part (b), $\|f_n\|_p = c_p$ independent of n and R, and so

$$2Nc_p = 2\sum_{n=1}^N \|f_n\|_p$$

and by applying triangle inequality through the sum, we have

$$\|S_n\|_p \leq \sum_{n=1}^N \|f_n\|_p = c_n \implies \|S_n\|_p^p = Nc_p^p \implies \|S_n\|_p = N^{1/p}c_p$$

combining everything, we have that

$$2Nc^{1/p} \geq ||S_N||_p \geq N^{1/p}c_p$$

which is what we desired to show.

(d) For $q < p$, the L_q norm of a sum of altered Gaussians would grow faster than what the previous inequality would permit due to non-overlapping supports in the Fourier domain. This implies that
$$\|\hat{f}\|_q \leq C_{p,q}\|f\|_p$$
cannot hold for $q < p$, which is a contradiction. Thus, the condition $q \geq p$ is required.

□

Problem 21 (Folland 9.4). *Suppose that U and V are open in $\mathbb{R}^n$ and $\Phi : V \to U$ is a C^∞ diffeomorphism. Explain how to define $F \circ \Phi$ if $\Phi \in D'(U)$ for any $F \in D'(V)$.*

Proof. Here, given $\phi \in \mathcal{D}$ we begin by defining the pullback of ϕ by a test function Φ by $\phi \circ \Phi^{-1}$, and the action of F defined on the pullback of a test function is given by $< F, \phi \circ \Phi^{-1} >$. From this definition, we have that

$$< F \circ \Phi, \phi >=< F, \phi \circ \Phi^{-1} >$$

by a change of variables within the integral. This is given by

$$< F \circ \Phi, \phi >= \int_{\mathbb{R}^n} f(\Phi^{-1}(x))\phi(x)|\det D\phi^{-1}(x)|dx$$

where $D\phi^{-1}(x)$ is the Jacobian of $\phi^{-1}(x)$. $< F \circ \Phi, \phi >\in \mathcal{D}'(\mathbb{R}^n)$ since $f(\Phi^{-1}(x))\phi(x)$ is continuous in the integral by the continuity of f and smoothness of Φ.

□

Problem 22 (Folland 9.5). *Suppose that f is continuously differentiable on $\mathbb{R}$ except at $x_1, \ldots, x_m$, where f has jump discontinuities, and that its pointwise derivative $\frac{df}{dx}$ (defined except at the x_j's) is in $L^1_{loc}(\mathbb{R})$. Then the distribution derivative f' of f is given by*

$$f' = \left(\frac{df}{dx}\right) + \sum_{j=1}^{m} \left[f(x_j^+) - f(x_j^-)\right] \tau_{x_j}\delta.$$

Proof. First, notice that $\frac{df}{dx}$ can be defined in the traditional sense away from the jump discontinuities $x_1, ..., x_m$. The distributional derivative f' is given by

$$< f', \phi >= -\int_{\mathbb{R}} f(x)\phi'(x)dx$$

where $\phi \in C_c^\infty(\mathbb{R})$, by an application of integration by parts. Now for each jump discontinuity x_i, we introduce the term $[f(x_i+) - f(x_i-)]\tau_{x_i}\delta$, where δ is the Dirac delta function centered at x_i. Now by the properties of the delta distribution, we have

$$< [f(x_i+) - f(x_i-)]\tau_{x_i}\delta, \phi >= [f(x_i+) - f(x_i-)]\phi$$
$$\implies \sum_{i=1}^{m} < [f(x_i+) - f(x_i-)]\tau_{x_i}\delta, \phi >= \sum_{i=1}^{m}[f(x_i+) - f(x_i-)]\phi$$

Now accounting for the classical derivative and jump discontinuities in the distribution of f, we have

$$< f', \phi >= - \int_{\mathbb{R}} f(x)\phi'(x)dx + \sum_{i=1}^{m} [f(x_i+) - f(x_i-)]\phi$$

and by integration by parts, we have

$$< f', \phi >= \int_{\mathbb{R}} (\frac{df}{dx})\phi dx + \sum_{i=1}^{m} [f(x_i+) - f(x_i-)]\phi$$

and as this holds for all test functions ϕ, we conclude that

$$f' = \frac{df}{dx} + \sum_{i=1}^{m} [f(x_i+) - f(x_i-)]\tau_{x_i}\delta$$

□

Problem 23. *Let $l_k(f) = k[f(\frac{1}{k}) - f(-\frac{1}{k})]$ be distributions on $\mathbb{R}$. Show that $l_k \to l$ in the sense of distributions to some $l \in \mathcal{D}^*$, and find l.*

Proof. Here it is clear that as $k \to \infty$, we would expect $f(\frac{1}{k}) - f(-\frac{1}{k}) \to 0$, and so we will consider a neighborhood around 0 of f. At $x = 0$, the Taylor expansion of f is given by

$$f(x) = f(0) + xf'(0) + O(x^2)$$

and now substituting $x = \pm\frac{1}{k}$, we have

$$f(\frac{1}{k}) - f(-\frac{1}{k}) = \frac{2}{k}f'(0) + O(\frac{1}{k^2})$$

multiplying by k yields

$$l_k(f) = k(\frac{2}{k}f'(0) + O(\frac{1}{k^2})) = 2f'(0) + O(\frac{1}{k})$$

and now taking $k \to \infty$, we have

$$\lim_{k\to\infty} l_k(f) = l(f) = 2f'(0)$$

and we can pair this result with the distributional derivative of the delta function to obtain

$$l(f) = 2f'(0) = 2 < \delta', f >$$

and thus $l_k \to l$ in the sense of distributions. □

Problem 24. *We say a distribution $l \in \mathcal{D}^*$ is positive if $l(u) \geq 0$ for all $u \in \mathcal{D}$ with $u \geq 0$. Classify all positive distributions [Hint: you actually did this on a past homework].*

Proof. Denote the space of positive distributions as

$$\mathcal{D}^*_{\mathcal{P}} = \{l \in D^*_P : l(u) \geq 0, \forall u \in \mathcal{D}\}$$

For an an arbitrary $T \in \mathcal{D}^*_{\mathcal{P}}$, T as a distribution (and positive linear functional) is uniquely characterized some Radon measure $\mu \in \mathbb{R}^n$ by the Riesz Representation Theorem, that is, for non-negative $\phi \in C_c^\infty(\mathbb{R}^n)$, we have

$$< T, \phi >= \int_{\mathbb{R}^n} \phi d\mu$$

Thus, for any non-negative $\phi \in C_c^\infty(\mathbb{R}^n)$, T as a positive distribution uniquely characterized as integration against a positive Radon measure. $\square$

Problem 25 (Folland 9.17)**.** *Suppose that $F \in S'$. Show that*

(a) $(\tau_y F)^\wedge = e^{-2\pi i\xi\cdot y}\hat{F}, \quad \tau_n F = [e^{2\pi in\cdot x}F]^\wedge.$

(b) $\partial^\alpha F = [(-2\pi ix)^\alpha F]^\wedge, \quad (\partial^\alpha F)^\wedge = (2\pi i\xi)^\alpha \hat{F}.$

(c) $(F \circ T)^\wedge = |\det T|^{-1} F \circ (T^*)^{-1}$ *for* $T \in GL(n, \mathbb{R})$.

(d) $(F * \psi) = \psi\hat{F}$ *for* $\psi \in S$.

Proof. (a) First, notice that using the properties of tempered distributions, we have that

$$< \widehat{\tau_y F}, \phi >=< \tau_y F, \widehat{\phi} >=< F, \tau_{-y}\widehat{\phi} >$$

and since $\tau_{-y}\widehat{\phi} = \widehat{\phi(x+y)}$, and by definition,

$$\widehat{\phi(x+y)} = \int_{\mathbb{R}^n} e^{-2\pi i(x+y)\xi}\phi\xi d\xi = e^{-2\pi iy\xi}\widehat{\phi(x)}$$

and thus

$$< F, \tau_{-y}\widehat{\phi} >=< F, e^{-2\pi iy\xi}\widehat{\phi(x)} >=< F, \widehat{\phi} > e^{-2\pi iy\xi} =< \widehat{F}, \phi > e^{-2\pi iy\xi}$$

and hence we have that $\widehat{\tau_{-y}F} = e^{-2\pi iy\xi}\widehat{F}$ as required.

(b) First of all, we know that for a test function in the Schwarz Space, we have that

$$< \partial^\alpha F, \phi >= (-1)^\alpha < F, \partial^\alpha \phi >$$

and note that

$$\partial^\alpha \hat{\phi} = (2\pi i)^{|\alpha|} \int e^{-2\pi ix\cdot s} x^\alpha \phi(x) dx$$

and thus we have that

$$< \widehat{\partial^\alpha F}, \phi >= (-1)^{|\alpha|} < F, \partial^\alpha \hat{\phi} >= (-1)^{|\alpha|}(2\pi i)^{|\alpha|} < F, x^\alpha \hat{\phi} >$$

and since $(-1)^{|\alpha|}(2\pi i)^{|\alpha|} = (2\pi i)^{|\alpha|}$, we have that

$$< \widehat{\partial^\alpha F}, \phi >= (2\pi is)^{|\alpha|} < \hat{F}, \phi > \implies \widehat{\partial^\alpha F(s)} = (2\pi is)^\alpha F\hat{(}s)$$

which is what we desired to show.

(c) First, note that $\widehat{(F \circ T)}(\phi) = f \circ T(\hat{\phi})$. Using a change of variables, we see that for $x \in \mathbb{R}^n$

$$\hat{\phi}(T^{-1}x) = |\det T|^{-1} \int_{\mathbb{R}^n} e^{-2\pi i (T^{-1}x, \xi)} \phi \xi d\xi$$

but we also have that $T * T^{-1}$ is the identity, and so

$$\hat{\phi}(T^{-1}x) = |\det T|^{-1} \int_{\mathbb{R}^n} e^{-2\pi i (x, (T*)^{-1}\xi} \phi(\xi) d\xi$$

and from this we obtain

$$\widehat{(F \circ T)}(\phi) = |\det T|^{-1} F(\widehat{\phi \circ (T*)^{-1}}) = |\det T|^{-1} \hat{F}((T*)^{-1}\xi)(\phi)$$

(d) First, note that $< F * \psi(x), \phi > = < F, \psi * \psi(-x) >$ and taking the Fourier transform, we have

$$\widehat{(F * \psi)}(\xi) = \hat{F}(\xi) \cdot \widehat{\psi * \phi(-x)} = \hat{F}(\xi) \cdot \hat{\psi} \cdot \phi(\hat{-}x)$$

since the Fourier transform of a convolution is a product of Fourier transforms. Substituting this back into the original equation, we have that

$$\widehat{(F * \psi)} = \hat{F}\hat{\psi}$$

as required.

$\square$

Problem 26 (Problem 10.5). *If X is a random variable with distribution $dP_X(t) = f(t)\,dt$ where $f(t) = f(-t)$, then the distribution of X^2 is $dP_{X^2}(t) = t^{-1/2} f(t^{1/2}) \chi_{(0,\infty)}(t)\,dt$*

Proof. Here, notice that for $t < 0$, the probability $F_{X^2}(t) = 0$ since X^2 is never negative. Now for $t \geq 0$, we have that $x^2 \leq t$ if X is in the interval $[-t^{1/2}, t^{1/2}]$ and hence

$$F_{X^2}(t) = P(Y \leq t) = P(-t^{1/2} \leq X \leq t^{1/2})$$

Now note that if $f(t)$ is the probability density function of X, then we have that

$$F_{X^2}(t) = \int_{-t^{1/2}}^{t^{1/2}} f(x) dx = 2 \int_0^{t^{1/2}} f(x) dx$$

by the symmetry of the integral. By differentiating, we have

$$\frac{d}{dt} F_{X^2}(t) = 2f(t^{1/2}) \frac{d}{dt}(t^{1/2}) = t^{-1/2} f(t^{1/2})$$

by the fundamental theorem of calculus. Now accounting for the fact that the probability density function is identically zero for $t < 0$, we multiply by the characteristic function on $(0, \infty)$ to obtain

$$dP_{X^2}(t) = t^{-1/2} f(t^{1/2}) \chi_{(0,\infty)}(t) dt$$

$\square$

Problem 27 (Folland 7.7)**.** *Let δ_t denote the point mass at $t \in \mathbb{R}$. Given $0 < p < 1$, let $\beta_p = p\delta_1 + (1-p)\delta_0$, and let β_p^{*n} be the nth convolution power of β_p. Then*

$$\beta_p^{*n} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \delta_k,$$

and the mean and variance of β_p^{*n} are np and $np(1-p)$, respectively. β_p^{*n} is called the binomial distribution on $\{0, \dots, n\}$ with parameter p.

Proof. Notice that the n-th convolution power β_p^{*n} is the measure corresponding to the sum of n independent variables each distributed according to β_p. We know that the probability mass function is given by $\binom{n}{k} p^k (1-p)^{n-k}$, and hence β_p^{*n} is the measure which picks out the probability mass function at each point k, which yields

$$\beta_p^{*n} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$$

From this, we obtain that the mean μ is given by

$$\mu = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

which simplifies to np by binomial theorem. Moreover, the variance σ^2 is given by

$$\sigma^2 = \sum_{k=0}^{n} (k-\mu)^2 \binom{n}{k} p^k (1-p)^{n-k}$$

which simplifies to $np(1-p)$ after recognizing that the variance of a Bernoulli random variable is $p(1-p)$, summed n times. □

Problem 28. *This is a modification of Folland 10.7 to avoid using convolutions explicitly.*

(a) Let $\Omega_n = \{0,1\}^n$, $\mathcal{F}_n = 2^{\Omega_n}$ be the entire power set. For a $p \in (0,1)$, construct a probability on this space as a product $P(E_1 \times \ldots \times E_n) = q(E_1)q(E_2)\ldots q(E_n)$, where $q(\{1\}) = p$, $q(\{0\}) = 1-p$ and then extending to all events by additivity. Check that this gives a well-defined probability.

(b) Let $X_k(\omega) = \omega_k$, i.e., X_k is 1 if the k-th entry is 1 and 0 if it is 0. Find the distribution μ_{X_k} and the expectation and variance of X_k.

(c) Let $S_k = \sum_{j=1}^{k} X_k$. Compute the expectation and variance of S_k.

(d) Find the distribution μ_{S_k}. [This is the formula in Folland 10.7.]

Proof. (a) First, it is clear that P is non-negative, as it is the product of finitely many $q(E_j)$ which are all non-negative since $0 < p < 1$. Moreover, we have that $P(\Omega_n) = q(\{0,1\}) \times \cdots \times q(\{0,1\})$. However, $q(\{0,1\}) = q(\{1\}) + q(\{0\}) = p + (1-p) = 1$ and hence $P(\Omega_n) = 1 \times \cdots \times 1 = 1$. P is also finitely additive due to the fact that all events in Ω_n are independent (and finite), fulfilling countable additivity as well.

(b) Since $X_k(\omega) = \omega_k$ can only take two values and each entry in ω is independent, this is a Bernoulli distribution with parameter p, and this distribution is given by $\mu_{X_k}(0) = 1-p$ and $\mu_{X_k}(1) = p$. The expectation is given by

$$\mathbb{E}[X_k] = 0(\mu_{X_k}(0)) + 1(\mu_{X_k}(1)) = p$$

and the variance is given by

$$\mathrm{Var}(X_k) = \mathbb{E}[(X_k)^2] - (\mathbb{E}[X_k])^2 = p - p^2 = p(1-p)$$

(c) Here, we see that S_k follows a binomial distribution with k and p. We have that

$$\mathbb{E}[S_k] = \mathbb{E}[\sum_{j=1}^{k} X_j] = \sum_{j=1}^{k} \mathbb{E}[X_j] = \sum_{j=1}^{k} p = kp$$

and moreover

$$\mathrm{Var}(S_k) = \sum_{j=1}^{k} \mathrm{Var}(X_j) = \sum_{j=1}^{k} p(1-p) = kp(1-p)$$

(d) Each X_j in the sum is a Bernoulli distribution which follows a binomial distribution for variables n and p. This corresponds to

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Now for the random variable S_k, there are k trials and the probability of success for each of them is p, giving:

$$P(S_k = x) = \binom{k}{x} p^x (1-p)^{k-x}$$

and in fact this is the binomial distribution μ_{S_k} for S_k.

□