Mathematics 416 | Final Exam Spring 2008 | Name: Dr. Grow

This is an open-textbook, open-notes test. That is, while solving the problems on this examination you may refer to your textbooks, Royden's *Real Analysis* and Karatzas and Shreve's *Brownian Motion and Stochastic Calculus*, or to the lecture notes you have taken for Math 416 this semester. Please solve any **three** of the following four problems and **circle** the number of each problem you want me to grade. Each problem is of equal value: 100 points. Thus, the total possible number of points on this examination is 300.

1. Let E denote a Borel subset of a separable metric space (X,d), let $m_\alpha(E)$ denote the Hausdorff α-dimensional measure of E, and let P denote the Cantor ternary set in the interval $[0,1]$.

25 (a) Show that $\inf\{\alpha : m_\alpha(E)=0\} = \sup\{\beta : m_\beta(E)=\infty\}$.

(Note: The common value of the preceding infimum and supremum is called the Hausdorff dimension of the set E.)

25 (b) If $\alpha > \log(2)/\log(3)$, show that $m_\alpha(P)=0$.

40 (c) If $\beta = \log(2)/\log(3)$, show that $m_\beta(P) = 1/2^\beta$.

10 (d) What is the Hausdorff dimension of P?

2. Let $f_1,\ldots,f_n$ be measurable real-valued functions on a probability space $(X,\mathcal{A},\mu)$. Prove or

$\Rightarrow)$ disprove: $\mathcal{F}=\{f_1,\ldots,f_n\}$ is an independent family if and only if the cumulative joint distribution

$\Leftarrow)$ 85 function for the ordered n-tuple $(f_1,\ldots,f_n)$ satisfies

$$F_{(f_1,\ldots,f_n)}(t_1,\ldots,t_n) = F_{f_1}(t_1)\ldots F_{f_n}(t_n)$$

for all $(t_1,\ldots,t_n)\in\mathbb{R}^n$.

50 3. (a) Let $f \in L^2(0,1)$ have the property that $\hat{f}(n) \neq 0$ for all integers n. Show that the set of all convolution products $f * g$, as g ranges over $L^2(0,1)$, is dense in $L^2(0,1)$.

50 (b) Let $f \in L^2(0,1)$ have the property that the set of all convolution products $f * g$, as g ranges over $L^2(0,1)$, is dense in $L^2(0,1)$. Show that $\hat{f}(n) \neq 0$ for all integers n.

50 4. (a) Show that there is no function $d \in L^1(\mathbb{R}^n)$ such that $d * f = f$ for all $f \in L^1(\mathbb{R}^n)$.

50 (b) Show that there is a sequence of functions $\{d_k\}_{k=1}^\infty \subseteq L^1(\mathbb{R}^n)$ such that $d_k * f \to f$ in the $L^1(\mathbb{R}^n)$-norm for each $f \in L^1(\mathbb{R}^n)$.

#1. (a) The desired conclusion, $\inf\{\alpha : m_\alpha(E)=0\} = \sup\{\beta : m_\beta(E)=\infty\}$, easily follows from (1) and (2) below.

(1) If $m_{\alpha_0}(E) < \infty$ then $m_\alpha(E) = 0$ for all $\alpha > \alpha_0$.

(2) If $0 < m_{\alpha_0}(E)$ then $m_\beta(E) = \infty$ for all $\beta < \alpha_0$.

<u>Proof of (1)</u>: Suppose $m_{\alpha_0}(E) < \infty$ and let $\alpha > \alpha_0$. Then to each sufficiently small $\varepsilon > 0$, there corresponds a covering of open balls $\{B_i\}_{i=1}^\infty$ of E with each radius $r_i \in (0,\varepsilon)$ and $\sum_{i=1}^\infty r_i^{\alpha_0} < m_{\alpha_0}(E) + 1$. Then $\sum_{i=1}^\infty r_i^\alpha = \sum_{i=1}^\infty r_i^{\alpha-\alpha_0} \cdot r_i^{\alpha_0} < \varepsilon^{\alpha-\alpha_0} \sum_{i=1}^\infty r_i^{\alpha_0} < \varepsilon^{\alpha-\alpha_0}(m_{\alpha_0}(E)+1)$, which tends to 0 as $\varepsilon \to 0^+$. Thus

$$m_\alpha(E) = \lim_{\varepsilon\to 0^+} \lambda_\alpha^{(\varepsilon)}(E) = \lim_{\varepsilon \to 0^+} \inf\left\{\sum_{i=1}^\infty r_i^\alpha : E \subseteq \bigcup_{i=1}^\infty B_i,\ \text{each } r_i \in (0,\varepsilon)\right\}$$
$$= 0.$$

<u>Proof of (2)</u>: Suppose $0 < m_{\alpha_0}(E)$ and let $\beta < \alpha_0$. Then for each sufficiently small $\varepsilon > 0$ and for every covering of E by open balls $\{B_i\}_{i=1}^\infty$ with each radius $r_i \in (0,\varepsilon)$, we have $\sum_{i=1}^\infty r_i^{\alpha_0} > \frac{m_{\alpha_0}(E)}{2} > 0$. Consequently,

$$\sum_{i=1}^\infty r_i^\beta = \sum_{i=1}^\infty r_i^{\beta-\alpha_0} \cdot r_i^{\alpha_0} > \varepsilon^{\beta-\alpha_0} \sum_{i=1}^\infty r_i^{\alpha_0} > \varepsilon^{\beta-\alpha_0} \cdot \frac{m_{\alpha_0}(E)}{2}$$

which tends to ∞ as $\varepsilon \to 0^+$. Therefore

$$m_\beta(E) = \lim_{\varepsilon \to 0^+} \lambda_\beta^{(\varepsilon)}(E) = \lim_{\varepsilon\to 0^+} \inf\left\{\sum_{i=1}^\infty r_i^\beta : E \subseteq \bigcup_{i=1}^\infty B_i,\ \text{each } r_i \in (0,\varepsilon)\right\} = \infty.$$

#1. (b) Suppose $\alpha > \frac{\log(2)}{\log(3)}$; i.e. suppose $3^\alpha > 2$. Given $\varepsilon > 0$, choose $k_0 \in \mathbb{N}$ such that $1/3^{k_0} < \varepsilon$. For all $k \geq k_0$, 2^k closed disjoint intervals of length $1/3^k$ cover P. The interiors of these intervals cover all but a finite number of points of P. Therefore

$$\lambda_\alpha^{(\varepsilon)}(P) \leq \sum_{i=1}^{2^k} r_i^\alpha = \sum_{i=1}^{2^k}\left(\frac{1}{2}\cdot\frac{1}{3^k}\right)^\alpha = 2^k\left(\frac{1}{2\cdot 3^k}\right)^\alpha = \frac{1}{2^\alpha}\left(\frac{2}{3^\alpha}\right)^k \to 0$$

as $k \to \infty$. Thus $\lambda_\alpha^{(\varepsilon)}(P) = 0$ for all $\varepsilon > 0$ so $m_\alpha(P) = \lim_{\varepsilon \to 0^+} \lambda_\alpha^{(\varepsilon)}(P) = 0$.

(c) Suppose $\beta = \frac{\log(2)}{\log(3)}$; i.e. suppose $3^\beta = 2$. Let $\varepsilon > 0$ and let $\{B_i\}_{i=1}^\infty$ be a sequence of open balls (i.e. open intervals) in $[0,1]$ such that $P \subseteq \bigcup_{i=1}^\infty B_i$ and $r_i = \text{radius of } B_i \in (0,\varepsilon)$ for all $i \geq 1$. Since P is compact there exists a finite subcover $B_{i_1}, \ldots, B_{i_n}$ of P. Recall that $P = \bigcap_{k=1}^\infty P_k$ where each P_k consists of the union of 2^k closed intervals of length $1/3^k$ obtained at the k^{th} stage in the construction of the Cantor set. Let $\delta > 0$ and let $(1+\delta)B_{i_j}$ denote the dilate by $1+\delta$ of the ball B_{i_j} about its center (i.e. midpoint). Choose $k \in \mathbb{N}$ sufficiently large so $1/3^k < \delta \cdot \min\{r_{i_j} : 1 \leq j \leq n\}$. It follows that $\frac{1}{3^k} + r_{i_j} < (1+\delta)r_{i_j}$ for all $1 \leq j \leq n$, so if one of the closed intervals I of length $1/3^k$ in the collection of 2^k intervals generating P_k has nonempty intersection with some ball B_{i_j}, then $I \subset (1+\delta)B_{i_j}$. It follows that $P_k \subseteq \bigcup_{j=1}^n (1+\delta)B_{i_j}$ so

$$\frac{1}{2^\beta} = 2^k\cdot\left(\frac{1}{2}\cdot\frac{1}{3^k}\right)^\beta = \sum_{i=1}^{2^k} \rho_i^\beta \leq \sum_{j=1}^n \left[(1+\delta)r_{i_j}\right]^\beta \leq (1+\delta)^\beta \sum_{i=1}^\infty r_i^\beta .$$

ρ_i = radius of the i^{th} interval (ball) in the collection generating P_k

Since $\delta > 0$ is arbitrary, $\frac{1}{2^\beta} \leq \sum_{i=1}^{\infty} r_i^\beta$. Hence

$$\frac{1}{2^\beta} \leq \inf \left\{ \sum_{i=1}^{\infty} r_i^\beta : P \subseteq \bigcup_{i=1}^{\infty} B_i, \text{ each } r_i < \varepsilon \right\} = \lambda_\beta^{(\varepsilon)}(P).$$

Since $\varepsilon > 0$ is arbitrary, $\frac{1}{2^\beta} \leq \lim_{\varepsilon \to 0^+} \lambda_\beta^{(\varepsilon)}(P) = m_\beta(P)$. On the other hand, the argument in part (a) shows that if $1/3^k < \varepsilon$ then

$$\lambda_\beta^{(\varepsilon)}(P) \leq \sum_{i=1}^{2^k} \rho_i^\beta = \sum_{i=1}^{2^k} \left(\frac{1}{2} \cdot \frac{1}{3^k} \right)^\beta = \frac{1}{2^\beta} \left(\frac{2}{3^\beta} \right)^k = \frac{1}{2^\beta}.$$

Hence $m_\beta(P) = \frac{1}{2^\beta}$.

(d) $\quad \dim_{Haus}(P) = \dfrac{\log(2)}{\log(3)}$.

#2. Let $\mathcal{F} = \{f_1, \ldots, f_n\}$ be a set of n measurable, real-valued functions on a probability space $(X, \mathcal{A}, \mu)$.

($\Rightarrow$) Suppose that $\mathcal{F}$ is an independent family. If $(t_1, \ldots, t_n) \in \mathbb{R}^n$, then

$$F_{(f_1,\ldots,f_n)}(t_1, \ldots, t_n) = \mu\left(\bigcap_{i=1}^{n} \{x \in X : f_i(x) \le t_i\} \right)$$

$$= \prod_{i=1}^{n} \mu(\{x \in X : f_i(x) \le t_i\}) \qquad \text{(by independence)}$$

$$= \prod_{i=1}^{n} F_{f_i}(t_i).$$

($\Leftarrow$) Suppose that the cumulative joint distribution function for $(f_1, \ldots, f_n)$ satisfies $F_{(f_1,\ldots,f_n)}(t_1, \ldots, t_n) \overset{(*)}{=} \prod_{i=1}^{n} F_{f_i}(t_i)$ for all $(t_1, \ldots, t_n) \in \mathbb{R}^n$.

Define the (Borel-)measurable transformation $T : X \to \mathbb{R}^n$ by

$$T(x) = (f_1(x), \ldots, f_n(x)) \qquad (x \in X).$$

Let $(t_1, \ldots, t_n) \in \mathbb{R}^n$ and $M_i = (-\infty, t_i]$ for $1 \le i \le n$. Then

$$(\mu T^{-1})(M_1 \times \cdots \times M_n) = \mu(T^{-1}[M_1 \times \cdots \times M_n])$$

$$= \mu(\{x \in X : T(x) \in M_1 \times \cdots \times M_n\})$$

$$= \mu\left(\bigcap_{i=1}^{n} \{x \in X : f_i(x) \le t_i\} \right)$$

10 pts.

$$= F_{(f_1,\ldots,f_n)}(t_1, \ldots, t_n).$$

On the other hand,

$$
\begin{aligned}
(\mu f_1^{-1} \times \dots \times \mu f_n^{-1})(M_1 \times \dots \times M_n) &= (\mu f_1^{-1})(M_1) \cdot \ldots \cdot (\mu f_n^{-1})(M_n) \\
&= \prod_{i=1}^{n} \mu(\{x \in \mathbb{X} : f_i(x) \le t_i\}) \\
&= \prod_{i=1}^{n} F_{f_i}(t_i) \,.
\end{aligned}
$$

10 pts.

It follows then from (*) that μT^{-1} and $\mu f_1^{-1} \times \dots \times \mu f_n^{-1}$ are Borel measures on $\mathbb{R}^n$ which agree on the class of subsets of the form $I_1 \times \dots \times I_n$ where each $I_i = (-\infty, b_i]$ for some $b_i \in \mathbb{R}$. Let $\mathcal{C}$ denote the class of intervals of $\mathbb{R}$ of the form $I = (-\infty, b]$ for some $b \in \mathbb{R}$. It is easy to see that $\mathcal{C}$ is a π-system:

10 pts.

10 pts.

$$(-\infty, b_1] \cap (-\infty, b_2] = (-\infty, b_1 \wedge b_2] \in \mathcal{C} \,.$$

Let $\mathcal{D}$ denote the collection of Borel sets $B_1 \subseteq \mathbb{R}$ such that for any n sets from $\mathcal{D}$, say $B_1, \dots, B_n$, we have the equality

10 pts.

$$\mu T^{-1}(B_1 \times \dots \times B_n) = (\mu f_1^{-1} \times \dots \times \mu f_n^{-1})(B_1 \times \dots \times B_n)$$

10 pts. <u>Claim</u>: $\mathcal{D}$ is a λ-system.

<u>Remark</u>: Clearly $\mathcal{C} \subseteq \mathcal{D}$. It will then follow from the claim and Dynkin's π-λ theorem that $\mathcal{D}$ contains the σ-algebra generated by $\mathcal{C}$. But clearly $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$, so it will follow that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{D}$. That is, for all Borel sets $B_1, \dots, B_n$ in $\mathbb{R}$ we have

10 pts.

$$\mu\left(\bigcap_{i=1}^{n}\{x\in X: f_i(x)\in B_i\}\right) = \mu T^{-1}(B_1\times\cdots\times B_n)$$
$$= (\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})(B_1\times\cdots\times B_n)$$
$$= \prod_{i=1}^{n}\mu(\{x\in X: f_i(x)\in B_i\}),$$

and thus $\mathcal{F}=\{f_1,\ldots,f_n\}$ is an independent family.

15 pts. <u>Proof of Claim</u>:

(i′) Let $B_1=\mathbb{R}$ and $B_2,\ldots,B_n\in\mathcal{D}$. Then for each $b\in\mathbb{R}$, $(-\infty,b]\in\mathcal{D}$ so

$$\mu T^{-1}((-\infty,b]\times B_2\times\cdots\times B_n) = (\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})((-\infty,b]\times B_2\times\cdots\times B_n).$$

Hence

$$\mu T^{-1}(\mathbb{R}\times B_2\times\cdots\times B_n) \overset{\#1,\ \text{Sec. 11.1}}{=} \lim_{m\to\infty}\mu T^{-1}((-\infty,m]\times B_2\times\cdots\times B_n)$$
$$= \lim_{m\to\infty}(\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})((-\infty,m]\times B_2\times\cdots\times B_n)$$
$$= (\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})(\mathbb{R}\times B_2\times\cdots\times B_n)$$

Since $B_2,\ldots,B_n\in\mathcal{D}$ were arbitrary, $B_1=\mathbb{R}\in\mathcal{D}$, by definition.

(ii′) Suppose $B_1\in\mathcal{D}$ and let $B_2,\ldots,B_n\in\mathcal{D}$. Then

$$\mu T^{-1}((\mathbb{R}\setminus B_1)\times B_2\times\cdots\times B_n) = \mu T^{-1}(\mathbb{R}\times B_2\times\cdots\times B_n) - \mu T^{-1}(B_1\times B_2\times\cdots\times B_n)$$
$$= (\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})(\mathbb{R}\times B_2\times\cdots\times B_n) - (\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})(B_1\times B_2\times\cdots\times B_n)$$
$$= (\mu f_1^{-1}\times\cdots\times\mu f_n^{-1})((\mathbb{R}\setminus B_1)\times B_2\times\cdots\times B_n)$$

Consequently $\mathbb{R}\setminus B_1\in\mathcal{D}$.

(iii′) Let $\{A_k\}_{k=1}^{\infty} \subseteq \mathcal{D}$ with $A_k \cap A_\ell = \phi$ if $k \neq \ell$, and let $B_2, \ldots, B_n \in \mathcal{D}$. Then

$$
\begin{aligned}
\mu T^{-1}\Big(\big(\dot{\bigcup_{k=1}^{\infty}} A_k\big) \times B_2 \times \cdots \times B_n \Big) &\overset{\#1,\ \text{Sec. II.1}}{=} \lim_{K\to\infty} \mu T^{-1}\Big(\big(\dot{\bigcup_{k=1}^{K}} A_k\big) \times B_2 \times \cdots \times B_n \Big) \\
&= \lim_{K\to\infty} \dot{\bigcup_{k=1}^{K}} (\mu T^{-1})(A_k \times B_2 \times \cdots \times B_n) \\
&= \lim_{K\to\infty} \dot{\bigcup_{k=1}^{K}} (\mu f_1^{-1} \times \cdots \times \mu f_n^{-1})(A_k \times B_2 \times \cdots \times B_n) \\
&= \lim_{K\to\infty} (\mu f_1^{-1} \times \cdots \times \mu f_n^{-1})\Big(\big(\dot{\bigcup_{k=1}^{K}} A_k\big) \times B_2 \times \cdots \times B_n \Big) \\
&= (\mu f_1^{-1} \times \cdots \times \mu f_n^{-1})\Big(\big(\dot{\bigcup_{k=1}^{\infty}} A_k\big) \times B_2 \times \cdots \times B_n \Big)
\end{aligned}
$$

Since $B_2, \ldots, B_n \in \mathcal{D}$ were arbitrary, the definition of $\mathcal{D}$ implies $\dot{\bigcup_{k=1}^{\infty}} A_k \in \mathcal{D}$. Therefore $\mathcal{D}$ is a λ-system. This concludes the proof of the claim and hence the solution of problem 2.

#3. Let $f \in L^2(0,1)$.

(a) Suppose that $\hat{f}(n) \neq 0$ for all $n \in \mathbb{Z}$. Let $h \in L^2(0,1)$ and let N be a positive integer. Then denoting $e_n(t) = e^{2\pi i n t}$ for $n \in \mathbb{Z}$, $t \in (0,1)$, we have $S_N(h;t) = \sum_{k=-N}^{N} \hat{h}(k) e_k(t) = \sum_{k=-N}^{N} \frac{\hat{h}(k)}{\hat{f}(k)} \cdot \hat{f}(k) e_k(t) = \sum_{k=-N}^{N} \frac{\hat{h}(k)}{\hat{f}(k)} (f * e_k)(t)$

20 pts. $= f * \left(\sum_{k=-N}^{N} \frac{\hat{h}(k)}{\hat{f}(k)} e_k \right)(t)$. Note that $g_N(t) = \sum_{k=-N}^{N} \frac{\hat{h}(k)}{\hat{f}(k)} e_k(t)$

10 pts. is a trigonometric polynomial and hence belongs to $L^2(0,1)$. Also $S_N(h;t) = (f * g_N)(t)$ for all $t \in (0,1)$ and $\| S_N(h) - h \|_{L^2(0,1)} \to 0$ 10 pts. as $N \to \infty$. (This is a consequence of Parseval's identity; see the Math 315 lecture notes on the L^2-theory of Fourier series or Rudin's "Principles of Mathematical Analysis, Theorems 8.16 and 11.40.) It follows that the

10 pts set of all convolution products $f * g$, as g ranges over $L^2(0,1)$, is dense in $L^2(0,1)$.

(b) Suppose that the set of all convolution products $f * g$, as g ranges over $L^2(0,1)$, is dense in $L^2(0,1)$. Suppose $\hat{f}(n_0) = 0$ 10 pts. for some $n_0 \in \mathbb{Z}$. Then $\widehat{f*g}(n_0) = \hat{f}(n_0)\hat{g}(n_0) = 0$ for all 10 pts. $g \in L^2(0,1)$. For all $g \in L^2(0,1)$, we have by Parseval's identity

$$\| e_{n_0} - f*g \|^2_{L^2(0,1)} = 1 + \sum_{n \neq n_0} |-\widehat{f*g}(n)|^2 \geq 1, \qquad \text{20 pts.}$$

so the set $f * g$, as g ranges over $L^2(0,1)$, is not dense in $L^2(0,1)$, a contra-

#4. (a) Suppose that there exists a function $d \in L^1(\mathbb{R}^n)$ such that $d * f = f$ for all $f \in L^1(\mathbb{R}^n)$. In particular, $d * d = d$. If $\xi \in \mathbb{R}^n$ and $f \in L^1(\mathbb{R}^n)$, the Fourier transform of f at ξ is defined by $\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi} dx$. It is easy to show that $\widehat{(f*g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$ for all $f, g \in L^1(\mathbb{R}^n)$ and $\xi \in \mathbb{R}^n$. Hence $d * d = d$ implies $(\hat{d}(\xi))^2 = \widehat{(d*d)}(\xi) = \hat{d}(\xi)$ for all $\xi \in \mathbb{R}^n$ and hence, for each $\xi \in \mathbb{R}^n$, $\hat{d}(\xi)$ is either 0 or 1. But $\xi \mapsto \hat{f}(\xi)$ is a continuous function on $\mathbb{R}^n$ and satisfies $|\hat{f}(\xi)| \to 0$ as $\|\xi\| \to \infty$ when $f \in L^1(\mathbb{R}^n)$. It follows that $\hat{d}(\xi) = 0$ for all $\xi \in \mathbb{R}^n$ or $\hat{d}(\xi) = 1$ for all $\xi \in \mathbb{R}^n$, and clearly the second alternative is impossible since $|\hat{d}(\xi)| \to 0$ as $\|\xi\| \to \infty$. But then $\hat{d}(\xi) = 0$ for all $\xi \in \mathbb{R}^n$ implies $d(x) = 0$ a.e. by the uniqueness theorem, so $f(x) = (d*f)(x) = 0$ a.e. for all $f \in L^1(\mathbb{R}^n)$, a clear absurdity. Therefore, there is no $d \in L^1(\mathbb{R}^n)$ such that $d * f = f$ for all $f \in L^1(\mathbb{R}^n)$.

(b) For $R > 0$ and $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$, define the Fejer kernel by

$$K_R(\theta) = \frac{1}{(2\pi R)^n} \int_0^R \cdots \int_0^R \left(\int_{-\rho_1}^{\rho_1} \cdots \int_{-\rho_n}^{\rho_n} e^{i(\theta_1\xi_1 + \dots + \theta_n\xi_n)} d\xi_n \cdots d\xi_1 \right) d\rho_1 \cdots d\rho_n$$

[This kernel can be motivated by looking at the C^1-means of the partial Fourier integrals of $f \in L^1(\mathbb{R}^n)$: $\sigma_R(f;x) = \frac{1}{R^n}\int_0^R \cdots \int_0^R S_{(\rho_1,\ldots,\rho_n)}(f;x)\,d\rho_1\cdots d\rho_n$

$= \int_{\mathbb{R}^n} f(y) K_R(x-y)\,dy = (f * K_R)(x)$.] One proves the identities

$$K_R(\theta) = \frac{1}{(2\pi R)^n}\int_0^R \cdots \int_0^R \frac{2\sin(\rho_1\theta_1)}{\theta_1} \cdot \ldots \cdot \frac{2\sin(\rho_n\theta_n)}{\theta_n}\,d\rho_1\cdots d\rho_n$$

$$= \left(\frac{2\sin^2(R\theta_1/2)}{\pi R\theta_1^2}\right) \cdot \ldots \cdot \left(\frac{2\sin^2(R\theta_n/2)}{\pi R\theta_n^2}\right)$$

and the properties below follow as in the $n=1$ dimensional case (see HW set #3, problem E):

(1) $\int_{\mathbb{R}^n} K_R(\theta)\,d\theta = 1$ for all $R > 0$.

(2) $K_R(\theta) \geq 0$ for all $R > 0$ and all $\theta \in \mathbb{R}^n$.

(3) $\lim_{R\to\infty} \int_{\|\theta\|>\delta} K_R(\theta)\,d\theta = 0$ for every $\delta > 0$.

<u>Claim</u>: The sequence of functions $\{K_k\}_{k=1}^{\infty} \subseteq L^1(\mathbb{R}^n)$ has the property that $K_k * f \to f$ in the $L^1(\mathbb{R}^n)$-norm for each $f \in L^1(\mathbb{R}^n)$.

Proof of Claim: Let $f \in L^1(\mathbb{R}^n)$ and let $\varepsilon > 0$. Denote $f_y(x) = f(x-y)$ for $x, y \in \mathbb{R}^n$ and observe that $\|f - f_y\|_{L^1(\mathbb{R}^n)} \to 0$ as $y \to 0$. (This is clear if f is a continuous function of compact support on $\mathbb{R}^n$, and the density of such functions in $L^1(\mathbb{R}^n)$ yields the result for general $f \in L^1(\mathbb{R}^n)$.) Therefore there is a $\delta > 0$ such that $\|f - f_y\|_{L^1(\mathbb{R}^n)} < \frac{\varepsilon}{2}$ for all $\|y\| < \delta$. Choose $R_0 > 0$ such that $2\|f\|_{L^1(\mathbb{R}^n)} \int_{\|y\| \geq \delta} K_R(\theta)\, d\theta < \frac{\varepsilon}{2}$ for all $R \geq R_0$ (cf. property (3) of the Fejer kernel). For $k \geq R_0$ we have

$$\int_{\mathbb{R}^n} |K_k * f(x) - f(x)|\, dx \overset{\text{Property (1)}}{=} \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} K_k(y) f(x-y)\, dy - \int_{\mathbb{R}^n} K_k(y) f(x)\, dy \right| dx$$

$$\overset{\text{Property (2)}}{\leq} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K_k(y) |f(x-y) - f(x)|\, dy\, dx$$

$$= \int_{\mathbb{R}^n} \left(\int_{\|y\| < \delta} K_k(y) |f(x-y) - f(x)|\, dy \right) dx + \int_{\mathbb{R}^n} \left(\int_{\|y\| \geq \delta} K_k(y) |f(x-y) - f(x)|\, dy \right) dx$$

$$\overset{\text{Tonelli}}{=} \int_{\|y\| < \delta} K_k(y) \left(\int_{\mathbb{R}^n} |f(x-y) - f(x)|\, dx \right) dy + \int_{\|y\| \geq \delta} K_k(y) \left(\int_{\mathbb{R}^n} |f(x-y) - f(x)|\, dx \right) dy$$

$$\leq \int_{\|y\| < \delta} K_k(y) \|f - f_y\|_{L^1(\mathbb{R}^n)}\, dy + \int_{\|y\| \geq \delta} K_k(y)\, 2\|f\|_{L^1(\mathbb{R}^n)}\, dy$$

$$< \frac{\varepsilon}{2} \int_{\|y\|<\delta} K_k(y)\,dy + 2\|f\|_{L^1(\mathbb{R}^n)} \cdot \int_{\|y\|\geq\delta} K_k(y)\,dy$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

$$= \varepsilon .$$

Q.E.D.