factors) on \( R^k = R^1 \times \cdots \times R^1 \). Fubini's theorem of course gives a practical way to calculate volumes:

**Example 18.4.** Let \( V_k \) be the volume of the sphere of radius 1 in \( R^k \); by Theorem 12.2, a sphere in \( R^k \) with radius \( r \) has volume \( r^k V_k \). Let \( A \) be the unit sphere in \( R^k \), let \( B = [(x_1, x_2): x_1^2 + x_2^2 \leq 1] \), and let \( C(x_1, x_2) = [(x_3, \ldots, x_k): \Sigma_{i=3}^k x_i^2 \leq 1 - x_1^2 - x_2^2] \). By Fubini's theorem,

\[
V_k = \int_A dx_1 \cdots dx_k = \int_B dx_1 dx_2 \int_{C(x_1, x_2)} dx_3 \cdots dx_k
\]

\[
= \int_B dx_1 dx_2 V_{k-2} (1 - x_1^2 - x_2^2)^{(k-2)/2}
\]

\[
= V_{k-2} \int_0^1 \int_0^{2\pi} (1 - \rho^2)^{(k-2)/2} \rho \, d\rho \, d\theta
\]

\[
= \pi V_{k-2} \int_0^1 t^{(k-2)/2} \, dt = \frac{2\pi V_{k-2}}{k}.
\]

If \( V_0 \) is taken as 1, this holds for \( k = 2 \) as well as for \( k \geq 3 \). Since \( V_1 = 2 \), it follows by induction that

\[
V_{2i-1} = \frac{2(2\pi)^{i-1}}{1 \cdot 3 \cdot 5 \cdots (2i - 1)}, \quad V_{2i} = \frac{(2\pi)^i}{2 \cdot 4 \cdots (2i)}
\]

for \( i = 1, 2, \ldots \). \( \blacksquare \)

**PROBLEMS**

18.1. Show by Theorem 18.1 that if \( A \times B \) is nonempty and lies in \( \mathcal{F} \times \mathcal{G} \), then \( A \in \mathcal{F} \) and \( B \in \mathcal{G} \).

18.2. Suppose that \( X = Y \) is uncountable and \( \mathcal{F} = \mathcal{G} \) consists of the countable and the co-countable sets. Show that the diagonal \( E = [(x, y): x = y] \) does not lie in \( \mathcal{F} \times \mathcal{G} \), even though \( [y: (x, y) \in E] \in \mathcal{G} \) and \( [x: (x, y) \in E] \in \mathcal{F} \) for all \( x \) and \( y \).

18.3. Let \( (X, \mathcal{F}, \mu) = (Y, \mathcal{G}, \nu) \) be the completion of \( (R^1, \mathcal{B}^1, \lambda) \). Show that \( (X \times Y, \mathcal{F} \times \mathcal{G}, \mu \times \nu) \) is not complete.

18.4. The assumption of \( \sigma \)-finiteness in Theorem 18.2 is essential: Let \( \mu \) be Lebesgue measure on the line, let \( \nu \) be counting measure on the line, and take \( E = [(x, y): x = y] \). Then (18.1) and (18.2) do not agree.
18.5. Suppose $Y$ is the real line and $\nu$ is Lebesgue measure. Suppose that, for each $x$ in $X$, $f(x, y)$ has with respect to $y$ a derivative $f'(x, y)$. Formally,

$$\int_a^b \left[ \int_X f'(x, \xi) \mu(dx) \right] d\xi = \int_X \left[ f(x, y) - f(x, a) \right] \mu(dx)$$

and hence $\int_X f(x, y) \mu(dx)$ has derivative $\int_X f'(x, y) \mu(dx)$. Use this idea to prove Theorem 16.8(ii).

18.6. Suppose that $f$ is nonnegative on a $\sigma$-finite measure space $(\Omega, \mathcal{F}, \mu)$. Show that

$$\int_0^\infty f d\mu = (\mu \times \lambda)[(\omega, y) \in \Omega \times R^1 : 0 \leq y \leq f(\omega)].$$

Prove that the set on the right is measurable. This gives the "area under the curve." Given the existence of $\mu \times \lambda$ on $\Omega \times R^1$, one can use the right side of this equation as an alternative definition of the integral.


18.8. Suppose that $\nu(y : (x, y) \in E) = \nu(y : (x, y) \in F)$ for all $x$ and show that $(\mu \times \nu)(E) = (\mu \times \nu)(F)$. This is a general version of Cavalieri’s principle.

18.9. (a) Suppose that $\mu$ is $\sigma$-finite and prove the corollary to Theorem 16.7 by Fubini’s theorem in the product of $(\Omega, \mathcal{F}, \mu)$ and $(1, 2, \ldots)$ with counting measure.

(b) Relate the series in Problem 17.9 to Fubini’s theorem.

18.10. (a) Let $\mu = \nu$ be counting measure on $X = Y = \{1, 2, \ldots\}$. If

$$f(x, y) = \begin{cases} 2 - 2^{-x} & \text{if } x = y, \\ -2 + 2^{-x} & \text{if } x = y + 1, \\ 0 & \text{otherwise,} \end{cases}$$

then the iterated integrals exist but are unequal. Why does this not contradict Fubini’s theorem?

(b) Show that $xy/(x^2 + y^2)^2$ is not integrable over the square $[(x, y) : |x|, |y| \leq 1]$ even though the iterated integrals exist and are equal.

18.11. For an integrable function $f$ on $(0, 1]$, consider a Riemann sum in the form $R = \sum_{j=1}^n f(x_j)(x_j - x_{j-1})$, where $0 = x_0 < \cdots < x_n = 1$. Extend $f$ to $(0, 2]$ by setting $f(x) = f(x - 1)$ for $1 < x \leq 2$, and define

$$R(t) = \sum_{j=1}^n f(x_j + t)(x_j - x_{j-1}).$$

For $0 < t < 1$, the points $x_j + t$ reduced modulo 1 give a partition of $(0, 1]$, and $R(t)$ is essentially the corresponding Riemann sum, differing from it by at most three terms. If $\max(x_j - x_{j-1})$ is small, then $R(t)$ is for most values of $t$ a good approximation to $\int_0^1 f(x) \, dx$, even though $R = R(0)$ may
be a poor one. Show in fact that

\begin{align*}
(18.17) \quad \int_0^1 \left| R(t) - \int_0^1 f(x) \, dx \right| \, dt \\
= \int_0^1 \left| \sum_{j=1}^n \int_{x_{j-1}}^{x_j} \left[ f(x + t) - f(x + t) \right] \, dx \right| \, dt \\
\leq \sum_{j=1}^n \int_{x_{j-1}}^{x_j} \, \int_0^1 \left| f(x + t) - f(x + t) \right| \, dt \, dx.
\end{align*}

Given \( \epsilon \) choose a continuous \( g \) such that \( \int_0^1 |f(x) - g(x)| \, dx < \epsilon^2/3 \) and then choose \( \delta \) so that \( |x - y| < \delta \) implies that \( |g(x) - g(y)| < \epsilon^2/3 \). Show that \( \max(x_j - x_{j-1}) < \delta \) implies that the first integral in (18.17) is at most \( \epsilon^2 \) and hence that \( |R(t) - \int_0^1 f(x) \, dx| < \epsilon \) on a set of \( t \) in \( (0,1) \) of Lebesgue measure exceeding \( 1 - \epsilon \).

18.12. Exhibit a case in which (18.12) fails because \( F \) and \( G \) have a common point of discontinuity.

18.13. Prove (18.16) for the case in which all the functions are continuous by differentiating with respect to the upper limit of integration.

18.14. Prove for distribution functions \( F \) that \( \int_{-\infty}^{\infty} (F(x + c) - F(x)) \, dx = c \).

18.15. Prove for continuous distribution functions that \( \int_{-\infty}^{\infty} F(x) \, dF(x) = \frac{1}{2} \).

18.16. Suppose that a number \( f_n \) is defined for each \( n \geq n_0 \) and put \( F(x) = \sum_{n \leq n \leq x} f_n \). Deduce from (18.15) that

\begin{align*}
(18.18) \quad \sum_{n_0 \leq n \leq x} G(n) f_n = F(x) G(x) - \int_{n_0}^{x} F(t) g(t) \, dt
\end{align*}

if \( G(y) - G(x) = \int_{x}^{y} g(t) \, dt \), which will hold if \( G \) has continuous derivative \( g \). First assume that the \( f_n \) are nonnegative.

18.17. \( \uparrow \) Take \( n_0 = 1 \), \( f_n = 1 \), and \( G(x) = 1/x \) and derive \( \sum_{n \leq x} n^{-1} = \log x + \gamma + O(1/x) \), where \( \gamma = 1 - \int_{[1]}^{\infty} (t - [t]) t^{-2} \, dt \) is Euler's constant.

18.18. 5.17 18.16 \( \uparrow \) Use (18.18) and (5.47) to prove that there exists a constant \( c \) such that

\begin{align*}
(18.19) \quad \sum_{p \leq x} \frac{1}{p} = \log \log x + c + O\left( \frac{1}{\log x} \right).
\end{align*}

18.19. If \( A_k = \sum_{i=1}^{k} a_i \) and \( B_k = \sum_{i=1}^{k} b_i \), then

\begin{align*}
\sum_{k=1}^{n} a_k B_k = A_n B_n + \sum_{k=2}^{n} A_{k-1} b_k.
\end{align*}

This is Abel's partial summation formula. Derive it by arguing as in the proof of Theorem 18.4.
18.20. (a) Let \( I_n = \int_0^{\pi/2} (\sin x)^n \, dx \). Show by partial integration that \( I_n = (n - 1)(I_{n-2} - I_n) \) and hence that \( I_n = (n - 1)n^{-1}I_{n-2} \).
(b) Show by induction that
\[
I_{2n+1}^{-1} = (-1)^n (2n + 1) \left( \frac{-\frac{1}{2}}{n} \right).
\]
From (17.17) deduce that
\[
\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots
\]
and hence that
\[
\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots.
\]
(c) Show by induction that
\[
I_{2n} = \frac{2n - 1}{2n} \frac{2n - 3}{2n - 2} \cdots \frac{1}{4^2} \frac{\pi}{2}, \quad I_{2n+1} = \frac{2n + 1}{2n + 2} \frac{2n - 2}{2n - 1} \cdots \frac{3}{5} \frac{4}{3} \frac{2}{1}.
\]
From \( I_{2n-1} > I_{2n} > I_{2n+1} \) deduce that
\[
\frac{\pi}{2} \frac{2n}{2n + 1} < \frac{2^2 \cdot 4^2 \cdots (2n)^2}{3^2 \cdot 5^2 \cdots (2n - 1)^2(2n + 1)} < \frac{\pi}{2},
\]
and from this derive Wallis's formula,
\[
\frac{\pi}{2} = \frac{2 \cdot 4 \cdot 6 \cdot 8 \cdots}{1 \cdot 3 \cdot 5 \cdot 7 \cdots}
\]

18.21. Stirling's formula. The area under the curve \( y = \log x \) between the abscissas 1 and \( n \) is \( A_n = n \log n - n + 1 \). The area under the corresponding inscribed polygon is \( B_n = \log n! - \frac{1}{2} \log n \).
(a) Let $S_k$ be the sliver-shaped region bounded by curve and polygon between the abscissas $k$ and $k + 1$, and let $T_k$ be this region translated so that its right-hand vertex is at $(2, \log 2)$. Show that the $T_k$ do not overlap and that $T_n \cup T_{n+1} \cup \cdots$ is contained in a triangle of area $\frac{1}{2} \log(1 + 1/n)$. Conclude that

$$\log n! = \left( n + \frac{1}{2} \right) \log n - n + c + \alpha_n,$$

where $c$ is 1 minus the area of $T_1 \cup T_2 \cup \cdots$ and

$$0 < \alpha_n < \frac{1}{2} \log \left( 1 + \frac{1}{n} \right) < \frac{1}{n}.$$

(b) By Wallis's formula (18.20) show that

$$\log \sqrt{\frac{\pi}{2}} = \lim_n \left[ 2n \log 2 + 2 \log n! - \log(2n)! - \frac{1}{2} \log(2n + 1) \right].$$

Substitute (18.21) and show that $c = \log \sqrt{2\pi}$. This gives Stirling's formula

$$n! \sim \sqrt{2\pi n^{n+\frac{1}{2}} e^{-n}}.$$

18.22. Euler's gamma function is defined for positive $t$ by $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} \, dx$.

(a) Prove that $\Gamma^{(k)}(t) = \int_0^\infty x^{t-1} (\log x)^k e^{-x} \, dx$.

(b) Show by partial integration that $\Gamma(t + 1) = t\Gamma(t)$ and hence that $\Gamma(n + 1) = n!$ for integral $n$.

(c) From (18.10) deduce $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(d) Show that the unit sphere in $R^k$ has volume (see Example 18.4)

$$V_k = \frac{\pi^{k/2}}{\Gamma\left( \frac{k}{2} + 1 \right)}.$$

18.23. By partial integration prove that $\int_0^\infty (\sin x)/x^2 \, dx = \pi/2$ and $\int_{-\infty}^\infty (1 - \cos x) x^{-2} \, dx = \pi$.

18.24. 17.15† Identifying $\pi$. If $A_\theta$ is the set of $(r \cos \varphi, r \sin \varphi)$ satisfying $r \leq 1$ and $0 \leq \varphi < \theta$, then by (17.21),

$$\lambda_2(A_\theta) = \frac{\pi \theta}{2 \pi}.$$

But by Fubini's theorem, if $0 < \theta < \pi/2$, then

$$\lambda_2(A_\theta) = \int_0^{\cos \theta} \frac{\sin \theta}{\cos \theta} x \, dx + \int_{\cos \theta}^1 (1 - x^2)^{1/2} \, dx
= \frac{1}{2} \sin \theta \cos \theta + \int_{\cos \theta}^1 (1 - x^2)^{1/2} \, dx.$$
Equate the right sides of (18.23) and (18.24), differentiate with respect to \( \theta \), and conclude that \( \pi_0 = \pi \). Thus the factor \( \pi_0 / \pi \) can be dropped in (17.22), which gives the formula (17.15) for integrating in polar coordinates.

18.25. Suppose that \( \mu \) is a probability measure on \((X, \mathcal{F})\) and that, for each \( x \) in \( X \), \( \nu_x \) is a probability measure on \((Y, \mathcal{G})\). Suppose further that, for each \( B \) in \( \mathcal{G} \), \( \nu_x(B) \) is, as a function of \( x \), measurable \( \mathcal{F} \). Regard the \( \mu(A) \) as initial probabilities and the \( \nu_x(B) \) as transition probabilities.

(a) Show that, if \( E \in \mathcal{F} \times \mathcal{G} \), then \( \nu_x[y: (x, y) \in E] \) is measurable \( \mathcal{F} \).

(b) Show that \( \pi(E) = \int_X \nu_x[y: (x, y) \in E] \mu(dx) \) defines a probability measure on \( \mathcal{F} \times \mathcal{G} \). If \( \nu_x = \nu \) does not depend on \( x \), this is just (18.1).

(c) Show that if \( f \) is measurable \( \mathcal{F} \times \mathcal{G} \) and nonnegative, then \( \int_Y f(x, y) \nu_x(dy) \) is measurable \( \mathcal{F} \). Show further that

\[
\int_{X \times Y} f(x, y) \pi(d(x, y)) = \int_X \left[ \int_Y f(x, y) \nu_x(dy) \right] \mu(dx),
\]

which extends Fubini's theorem (in the probability case). Consider also \( f \)'s that may be negative.

(d) Let \( \nu(B) = \int_X \nu_x(B) \mu(dx) \). Show that \( \pi(X \times B) = \nu(B) \) and

\[
\int_Y f(y) \nu(dy) = \int_X \left[ \int_Y f(y) \nu_x(dy) \right] \mu(dx).
\]

SECTION 19. HAUSDORFF MEASURE*

The theory of Lebesgue measure gives the area of a figure in the plane and the volume of a solid in 3-space, but how can one define the area of a curved surface in 3-space? This section is devoted to such geometric questions.

The Definition

Suppose that \( A \) is a set in \( \mathbb{R}^k \). For positive \( m \) and \( \epsilon \) put

\[(19.1)\quad h_{m, \epsilon}(A) = \inf c_m \Sigma \left( \text{diam } B_n \right)^m,
\]

where the infimum extends over countable coverings of \( A \) by sets \( B_n \) with diameters \( \text{diam } B_n = \sup \{|x - y|: x, y \in B_n\} \) less than \( \epsilon \). Here \( c_m \) is a positive constant to be assigned later. As \( \epsilon \) decreases, the infimum extends over smaller classes, and so \( h_{m, \epsilon}(A) \) does not decrease. Thus \( h_{m, \epsilon}(A) \) has a

*This section should be omitted on a first reading.