STUDY GUIDE FOR MATH 21a

What follows is a rough, section by section outline of the material that we covered this semester in Math 21a. The outline for the sections in Chapter 5 of Ostebee and Zorn is only relevant for the Regular and Physics sections; the outline for Chapters 2-5 of the Rosner text is only relevant for the BioChem sections; and the outline for the CS section is, of course, only relevant for the people in it.

The final exam is on Tuesday, January 22 at 9:15. Students with last names starting with A-Foley go to Science Center Lecture Hall A, and students with last names starting Fr-Z go to Science Center Lecture Hall B.

Chapter 1.1:

- R^3 = 3-d space, R^2 = 2-d space.
- Abstract spaces versus coordinates.
- Drawing and labeling axis. Rotating axis.
- Distances and midpoints.
- Graphing equations.
- Simple surfaces, spheres, ellipses, parabolas, saddles.
- Orientations.
- Graphing in 2-d. Graphs as curves.
- Curves which are not graphs.

Chapter 1.2 and Appendix A:

- Curves in the plane as graphs of functions and pieces of graphs.
- Parametrized curves in the plane.
- Parametrized curves in R^3. $t \rightarrow (x(t), y(t), z(t))$ for t ranging over some part of R.
- Explain the difference between a curve and its parametrization. The curve sits as a path in R^3; the parametrization tells the speed of an ant walking the path.
- Parametrization and reversing direction on the path.
- Polar coordinates and parametrized curves in polar coordinates.

Chapter 1.3:

- Vectors in the plane as arrows, or algebraically as 2-tuples (a, b). Vectors in space as arrows or 3-tuples (a, b, c).

- If a vector $\mathbf{v} = (a, b)$, then a and b are called the <u>components</u> of $\mathbf{v}$. Likewise for a 3-d vector, $\mathbf{v} = (a, b, c)$.
- A vector, $\mathbf{v}$, has a <u>length</u> or <u>magnitude</u>, $|\mathbf{v}| = (a^2 + b^2 + c^2)^{1/2}$.
- Adding two vectors, $\mathbf{v} + \mathbf{w}$. Multiplying by numbers (scalars) r $\mathbf{v}$ where r = 2, -.556, etc. Give graphical picture of addition and multiplication. (Also, multiplying by –1 to get $-\mathbf{v}$.)
- The zero vector $\mathbf{0} = (0, 0)$ or $(0, 0, 0)$ in 3-d.
- The triangle inequality $|\mathbf{v} + \mathbf{w}| \le |\mathbf{v}| + |\mathbf{w}|$.
- The <u>standard basis</u>, $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$. Thus, any $\mathbf{v} = (a, b, c)$ can be <u>decomposed</u> as $\mathbf{v} = a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k}$ in terms of the standard basis. (We shall play with other basis as well later on.)

Chapter 1.4:

- Vector valued functions of time, $t \to \mathbf{v}(t) = (a(t), b(t), c(t))$. Thus, as t changes, the end point of $\mathbf{v}(t)$ traces out a parameterized curve in R^3.
- For example, the line segment running between (x_0, y_0) and $(x_0 + a, y_0 + b)$ can be parametrized as $(x_0, y_0) + t\,(a, b)$ as t runs between 0 and 1 (written $t \in [0, 1]$). Writing $\mathbf{P}_0 = (x_0, y_0)$ and $\mathbf{v} = (a, b)$, this becomes $\mathbf{P}(t) = \mathbf{P}_0 + t\,\mathbf{v}$.
- The derivative of a vector valued function of time $\mathbf{v}(t)$ gives a new vector valued function of time, $\dfrac{d}{dt}\mathbf{v}(t) \equiv \mathbf{v}'(t) = (a'(t), b'(t), c'(t))$.
- The derivative $\mathbf{v}'(t)$ lives on a "different" version of R^3 as does $\mathbf{v}(t)$ since its tail sits naturally at $\mathbf{v}(t)$, not at the origin. This is because $\mathbf{v}'(t) = \lim_{\Delta t \to 0} (\mathbf{v}(t + \Delta t) - \mathbf{v}(t))/\Delta t$ and so is a vector which might naturally be interpreted as having its end, not at the origin, but at $\mathbf{v}(t)$.
- Lets use different notation: If we view $t \to \mathbf{r}(t)$ as a vector valued function which traces out a path in space, then $\dfrac{d}{dt}\mathbf{r} \equiv \mathbf{r}'(t)$ is the instantaneous velocity at time t of the particle on the path; $|\mathbf{r}'(t)|$ is its instantaneous speed, and $\dfrac{d}{dt}\mathbf{r}'(t) = \mathbf{r}''(t)$ is its instantaneous acceleration of the particle.
- Algebraic rules: $(\mathbf{v}(t) + \mathbf{w}(t))' = \mathbf{v}'(t) + \mathbf{w}'(t)$ and $(r(t)\,\mathbf{v}(t))' = r'(t)\,\mathbf{v}(t) + r(t)\,\mathbf{v}(t)$ when r is a function of t as well as $\mathbf{v}$.
- Anti-derivatives of $\mathbf{v}(t)$. This is the vector $\mathbf{f}(t) = \displaystyle\int_0^t \mathbf{v}(\tau)\,d\tau$. If $\mathbf{v}(t) = (a(t), b(t), c(t))$, then the vector $\mathbf{f}$ has components $(\displaystyle\int_0^t a(\tau)\,d\tau, \int_0^t b(\tau)\,d\tau, \int_0^t c(\tau)\,d\tau)$.
- Thus, if $\mathbf{v}(t) = \mathbf{r}'(t)$, then $\displaystyle\int_0^t \mathbf{v}(\tau)\,d\tau = \mathbf{r}(t) - \mathbf{r}(0)$.

Chapter 1.5:

- Newton's law: $m\,\mathbf{r}''(t) = \mathbf{F}$.
- Arc length: The distance traveled along a path $t \to \mathbf{r}(t)$ for $t_0 < t < t_1$ is $\int_{t_0}^{t_1} |\mathbf{r}'(\tau)|\,d\tau$.

Chapter 1.6:

- Dot product: If $\mathbf{v} = (a, b, c)$ and $\mathbf{w} = (e, f, g)$, then $\mathbf{v} \cdot \mathbf{w} = ae + bf + cg$. This is a multiplication rule which takes two vectors and gives a number. For two component vectors, $\mathbf{v} = (a, b)$ and $\mathbf{w} = (e, f)$, then $\mathbf{v} \cdot \mathbf{w} = ae + bf$.
- Length $|\mathbf{v}| = (\mathbf{v} \cdot \mathbf{v})^{1/2}$.
- Algebra of dot product: $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$ and $(r\mathbf{v}) \cdot \mathbf{w} = r(\mathbf{v} \cdot \mathbf{w})$.
- Unit vectors in the plane: $\mathbf{u} = (a, b)$ with $|\mathbf{u}| = 1$. Can be written as $\mathbf{u} = (\cos\theta, \sin\theta)$. Unit vectors in R^3 points in a specified direction. Unit vector in direction $\mathbf{v}$ is $\mathbf{u} = \mathbf{v}/|\mathbf{v}|$.
- Inner product in terms of angles: $\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\,|\mathbf{w}|\cos\Delta$ where Δ = angle between $\mathbf{v}$ and $\mathbf{w}$.
- In 3-d, the same applies: $\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\,|\mathbf{w}|\cos\Delta$.
- Orthogonality: $\mathbf{v}$ and $\mathbf{w}$ are orthogonal (point at right angles) if and only if $\mathbf{v} \cdot \mathbf{w} = 0$.
- Projection of $\mathbf{v}$ in direction of unit vector $\mathbf{u}$: This is the vector $(\mathbf{v} \cdot \mathbf{u})\,\mathbf{u}$. Projections along $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ of $\mathbf{v}$ give $\mathbf{v}$'s components in these directions.
- Derivatives and dot product: $(\mathbf{v} \cdot \mathbf{w})' = \mathbf{v}' \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}'$. Here, $\mathbf{v}$, $\mathbf{w}$ are vector functions of t. Thus, motion on the sphere or the circle in the plane has $\mathbf{v} \cdot \mathbf{v}' = 0$.

Chapter 1.7:

- Equations for lines and planes: The plane (call it $\prod$) through a point $\mathbf{P} = (x_0, y_0, z_0)$ with normal direction $\mathbf{n} = (a, b, c)$ $\prod = \{\mathbf{r} = (x, y, z): (\mathbf{r} - \mathbf{P}) \cdot \mathbf{n} = 0\} = \{\mathbf{r}: \mathbf{r} \cdot \mathbf{n} = \mathbf{P} \cdot \mathbf{n}\}$ = Those points $\mathbf{r}$ such that the equality $a\,x + b\,y + c\,z = a\,x_0 + b\,y_0 + c\,z_0$.
- Parametric equation for a line: $t \to \mathbf{P} + t\,\mathbf{v}$. Parametric equation for a plane requires two variables (think of the x-y plane, for instance): $(t, s) \to \mathbf{P} + t\,\mathbf{v} + s\,\mathbf{w}$ where $\mathbf{v}$ and $\mathbf{w}$ are any two non-colinear vectors (can be orthogonal) with $\mathbf{v} \cdot \mathbf{n} = \mathbf{w} \cdot \mathbf{n} = 0$.
- Non-parametric equation for a line: Choose two non-colinear vectors, $\mathbf{n}_1$ and $\mathbf{n}_2$ which are orthgonal (at $90°$) to $\mathbf{v}$. Then $L = \{\mathbf{r} = (x, y, z): \mathbf{r} \cdot \mathbf{n}_1 = \mathbf{P} \cdot \mathbf{n}_1 \ and \ \mathbf{r} \cdot \mathbf{n}_2 = \mathbf{P} \cdot \mathbf{n}_2\}$.
- Square of distance from a point $\mathbf{Q}$ to the line $L = \{\mathbf{P} + t\,\mathbf{v}\}$ is $d^2 = |\mathbf{Q} - \mathbf{P}|^2 - ((\mathbf{Q} - \mathbf{P}) \cdot \mathbf{v})^2/|\mathbf{v}|^2$.
- Distance from point $\mathbf{Q}$ to plane $\prod = \{\mathbf{r}: \mathbf{r} \cdot \mathbf{n} = \mathbf{P} \cdot \mathbf{n}\}$ is $d = |(\mathbf{Q} - \mathbf{P}) \cdot \mathbf{n}|$.

- Parametric equation for a line: $t \rightarrow \mathbf{P} + t\,\mathbf{v}$. Parametric equation for a plane requires two variables (think of the x-y plane, for instance): $(t, s) \rightarrow \mathbf{P} + t\,\mathbf{v} + s\,\mathbf{w}$ where $\mathbf{v}$ and $\mathbf{w}$ are any two non-colinear vectors (can be orthogonal) with $\mathbf{v} \bullet \mathbf{n} = \mathbf{w} \bullet \mathbf{n} = 0$.
- Differentiation and products: $(\mathbf{v} \bullet \mathbf{w})' = \mathbf{v}' \bullet \mathbf{w} + \mathbf{v} \bullet \mathbf{w}'$.

Chapter 1.8:

- Cross product: $\mathbf{v} = (a, b, c)$ and $\mathbf{w} = (e, f, g)$, then $\mathbf{v} \times \mathbf{w}$ is a new vector, perpendicular to both $\mathbf{v}$ and $\mathbf{w}$ with length $|\mathbf{v} \times \mathbf{w}| = |\mathbf{v}|\,|\mathbf{w}|\,|\sin(\Delta)|$ where Δ is the angle between $\mathbf{v}$ and $\mathbf{w}$.
- Right hand rule for determining the direction between $\mathbf{v}$ and $\mathbf{w}$.
- $\mathbf{v} \times \mathbf{w} = (bg - cf, ce - ag, af - be)$. Thus, $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$, $\mathbf{k} \times \mathbf{i} = \mathbf{j}$. Also, $\mathbf{v} \times \mathbf{v} = \mathbf{0}$ for any vector $\mathbf{v}$.
- Algebra: $\mathbf{v} \times \mathbf{w} = -\,\mathbf{w} \times \mathbf{v}$ and $(r\,\mathbf{v}) \times \mathbf{w} = r\,(\mathbf{v} \times \mathbf{w})$ and $\mathbf{v} \times (\mathbf{w} + \mathbf{u}) = \mathbf{v} \times \mathbf{w} + \mathbf{v} \times \mathbf{u}$.
- In particular, $\mathbf{v} \times \mathbf{w} = 0$ if and only if $\mathbf{w}$ is parallel to $\mathbf{v}$. (Compare with dot product.)
- A unit length normal vector to the plane spanned by $\mathbf{v}$ and $\mathbf{w}$: $\mathbf{n} = |\mathbf{v} \times \mathbf{w}|^{-1}\,(\mathbf{v} \times \mathbf{w})$.
- Formula for a line through $\mathbf{P} = (x_0, y_0, z_0)$ in direction $\mathbf{v}$: l consist of those $\mathbf{r} = (x, y, z)$ with the property that $\mathbf{v} \times (\mathbf{r} - \mathbf{P}) = 0$.
- Distance to the line l from a point $\mathbf{Q}$: $d = |\mathbf{v}|^{-1}\,|\mathbf{v} \times (\mathbf{Q} - \mathbf{P})|$.
- Differentiation: $\mathbf{v}(t)$ and $\mathbf{w}(t)$ then $(\mathbf{v} \times \mathbf{w})' = \mathbf{v}' \times \mathbf{w} + \mathbf{v} \times \mathbf{w}'$. Also, $(\mathbf{v} \bullet \mathbf{w})' = \mathbf{v}' \bullet \mathbf{w} + \mathbf{v} \bullet \mathbf{w}'$.
- Planar motion: $\mathbf{v} \times \mathbf{v}' = \mathbf{a} = $ constant vector. Motion on the sphere: $\mathbf{v} \bullet \mathbf{v}' = a = $ constant number.

Appendix A:

- Matrices: Arrays of numbers where the position in the array is important. Example: $m \times n$ matrix has m rows and n columns. An $m \times m$ matrix is a <u>square</u> matrix. These generalize vectors, which have 1 row and n columns.
- Index notation: M_{ij} denotes the entry in the i'th row and j'th column
- Adding two matrices of the same type: $(A + B)_{ij} = A_{ij} + B_{ij}$. $(A - B)_{ij} = A_{ij} - B_{ij}$.
- Multiplying a matrix A by a real number r: $(r\,A)_{ij} = r\,A_{ij}$.
- Multiplying an $m \times n$ matrix times an $n \times k$ matrix to get an $m \times k$ matrix. $(A\,B)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + A_{i3}B_{3j} + \cdots + A_{in}B_{nj}$.
- Note that A B is not defined unless B has the same number of rows as A has columns. Thus, unless A and B are square, $m \times m$ matrices, one or both of A B and B A is not defined.
- Suppose A and B are $m \times m$ matrices. Then $A\,B \neq B\,A$ in general!

- Symmetric matrices: (needs to be square matrix) $A_{ij} = A_{ji}$. Anti-symmetric: $A_{ij} = -A_{ji}$.
- The zero matrix $0_{ij} = 0$.
- The identity $m \times m$ matrix $I_{ij} = 1$ if $i = j$ and zero otherwise. Thus, $I A = A I = A$ for any other $m \times m$ matrix A. An inverse, A^{-1}, of an $m \times m$ matrix A is one such that $A^{-1} A = A A^{-1} = I$. Not all matrices have inverses!
- A diagonal $m \times m$ matrix: $D_{ij} = 0$ unless $i = j$. Only non-zero entries are on the diagonal.
- Trace of an $m \times m$ matrix: $Tr(A) = A_{11} + A_{22} + \cdots + A_{mm}$.
- Determinant of a 2×2 matrix: $Det(A) = A_{11}A_{22} - A_{12}A_{21}$.
- The Determinant of a 3×3 matrix: $Det(A) = A_{11}A_{22}A_{33} - A_{11}A_{23}A_{32} + A_{12}A_{23}A_{31} - A_{12}A_{33}A_{21} + A_{13}A_{21}A_{32} - A_{13}A_{22}A_{31}$.
- Cross products & determinants: $\mathbf{v} \times \mathbf{w} = \det \begin{pmatrix} i & j & k \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$.

Chapter 2.1:

- Functions $f(x, y)$ of 2-variables and $f(x, y, z)$ of three. E.g $f = x^2 + y^2 + z^2$, or $f = x - y^2 z$ or even $f = e^{xyz}$.
- Arrow notation, $f: R^3 \to R$. Domain and Range.
- Ways to represent functions of 2-variables: Graphs and level sets of functions of 2-variables. By definition, f is constant on each level set. Contour maps.
- Representing functions of 3-variables: Level surfaces for functions of 3-variables, the surfaces where $f(x, y, z) = $ constant.
- Planes and 'linear functions', $f(x, y, z) = ax + by + cz$.
- Vector valued functions.

Chapter 2.2:

- Partial derivatives: $f_x = \lim_{\Delta \to 0} (f(x+\Delta, y) - f(x, y)/\Delta$ and $f_y = \lim_{\Delta \to 0} (f(x, y+\Delta) - f(x, y)/\Delta$. This is just 1-variable calculus in each direction.
- Each of f_x and f_y is a new function. Gradient: $\nabla f = (f_x, f_y)$ is a vector function of x,y which keeps track of the two derivatives. In 3-d, $\nabla f = (f_x, f_y, f_z)$.
- $|\nabla f|$ measures absolute size of change of f; direction of ∇f gives direction of greatest change in f as can be seen for a linear function, $f = ax + by + cz$ where $\nabla f = (a, b, c)$ is normal to the level sets.
- $\nabla f = 0$ at extreme points, local max, min and also at saddles.

Chapter 2.3:

- Taylor's approximation in 1-variable of h(x): $L(x) = h(x_0) + h'(x_0) \cdot (x - x_0)$ is the best approximation of a function h by a 'linear' function near $x = x_0$. Here, L gets the value correct at x_0 and also the slope. Note that L depends on x_0.
- Taylor's approximation for 2, 3 variable f(x): $L(\mathbf{r}) = f(\mathbf{r}_0) + f_x(\mathbf{r}_0)(x - x_0) + f_y(\mathbf{r}_0)(y - y_0) + f_z(\mathbf{r}_0)(z - z_0)$ is the simplest approximation to f which gets the value write at $\mathbf{r} = \mathbf{r}_0$ and also gets all of the partial derivatives right at $\mathbf{r} = \mathbf{r}_0$. Note that L depends on $\mathbf{r}_0$. Note also that L can be written as $L = f(\mathbf{r}_0) + \nabla f(\mathbf{r}_0) \bullet (\mathbf{r} - \mathbf{r}_0)$.
- Thus, the level sets of L are planes. In particular, the level set where $L = f(\mathbf{r}_0)$ is a plane—it is called the 'tangent' plane to the level set where $f = f(\mathbf{r}_0)$ at the point $\mathbf{r} = \mathbf{r}_0$. This level set of L is, in a sense, the best approximation to the surface where $f = f(\mathbf{r}_0)$ at $\mathbf{r}_0$ by a plane.
- Note that $\nabla f(\mathbf{r}_0)$ points normal to the plane where $L = f(\mathbf{r}_0)$ and as L approximates the function f near $\mathbf{r}_0$, it may not surprise you that ∇f points normal to the level sets of f.

Chapter 2.4:

- $|\nabla f|$ measures absolute size of change of f; direction of ∇f gives direction of greatest change in f as can be seen for a linear function, $f = ax + by + cz$ where $\nabla f = (a, b, c)$ is normal to the level sets. Also, for $f = x^2 + y^2 + z^2$, $\nabla f = (2x, 2y, 2z)$ points radially outward as it should to point in the direction of fastest increase.
- Gradients and the linear approximation: Given some f(x, y, z), and a point $\mathbf{r}_0 = (x_0, y_0, z_0)$, the simplest function, L(x, y, z) that equals f at $\mathbf{r}_0$ and has the same partial derivatives as f does at $\mathbf{r}_0$ is $L(x, y, z) = f(\mathbf{r}_0) + f_x(\mathbf{r}_0)(x - x_0) + f_y(\mathbf{r}_0)(y - y_0) + f_z(\mathbf{r}_0)(z - z_0)$ which can be written in shorthand as $L(\mathbf{r}) = f(x_0) + (\nabla f)|_{\mathbf{r}=\mathbf{r}_0} \bullet (\mathbf{r} - \mathbf{r}_0)$ where $\mathbf{r}$ is shorthand for the vector (x, y, z).
- Note that L typically will differ if you change $\mathbf{r}_0$.
- Level set of f through $\mathbf{r}_0$ is approximated by the level set of L through $\mathbf{r}_0$. The latter is a plane, called the <u>tangent</u> plane to the level set at $\mathbf{r}_0$. It is the best approximation at $\mathbf{r}_0$ to this level set by a plane.
- Directional derivative of f in the direction of the unit vector $\mathbf{u}$ is defined to be the function whose value at $\mathbf{r} = (x, y, z)$ is equal to $\mathbf{u} \bullet \nabla f = |\nabla f| \cos\theta$. Here, θ is the angle between $\mathbf{u}$ and ∇f at $\mathbf{r}$. This is the same as $\lim_{\Delta \to 0} (f(\mathbf{r} + \Delta \mathbf{u}) - f(\mathbf{r}))/\Delta$, as can be seen by approximating f at $\mathbf{r}$ by the linear approximation L.

- The directional derivative computes the rate of change of f in the direction **u**. Thus, if you were moving along the line in the direction of **u** at unit speed, you would see the derivative of f at your point equal to **u**•∇f at the point.

Chapter 2.5:

- Note: There are functions which are not differentiable at various points (some at all points). However, except at the end of the course, and except for some special cases, like $f(\mathbf{r}) = 1/|\mathbf{r}|$ and $f(\mathbf{r}) = \ln(|\mathbf{r}|)$, you can safely assume that all functions that you meet are continuous, differentiable and have differentiable derivatives, and so on.

Chapter 2.6:

- Second derivatives, consider f(x, y). There are two derivatives, f_x and f_y, four second derivatives, f_{xx}, f_{xy}, f_{yx}, f_{yy}, eight third derivatives, etc. However, things are less complicated than might appear as: $f_{xy} = f_{yx}$. In general, taking a number m derivatives in the x-direction and n in the y direction of a function f is insensitive to the order in which these derivatives are taken. Thus, $f_{xxyxxy} = f_{yxxxxy} = f_{yyxxxx} = \cdots$ etc. This is true for f(x, y, z) also: $f_{xy} = f_{yx}$, $f_{xz} = f_{zx}$, $f_{yz} = f_{zy}$, and likewise order is immaterial for third and higher order derivatives as well.
- Some times, it is convenient to group the second derivatives into a matrix. For example if f is a function of 2-variables, f(x, y), then we write the first derivatives as a vector, $\nabla f = (f_x, f_y)$ and

  the second derivatives as a matrix, $f'' = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$. Note that it is symmetric as

  $f_{xy} = f_{yx}$.
- For a function of x, y, z, the matrix of second derivatives is 3×3: $f'' = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix}$.

- Second order Taylor's approximation: The best approximation, Q(**r**), near a point $\mathbf{r}_0$ of a function f(**r**) which gets $Q(\mathbf{r_0}) = f(\mathbf{r_0})$, and has the same first and second derivatives at $\mathbf{r}_0$ as does f is $Q(\mathbf{r}) = f(\mathbf{r_0}) + (\nabla f)(\mathbf{r_0})\bullet(\mathbf{r} - \mathbf{r_0}) + 2^{-1} (\mathbf{r} - \mathbf{r_0})\bullet(f''(\mathbf{r_0})\cdot(\mathbf{r} - \mathbf{r_0}))$. Here, $f''(\mathbf{r_0})$ is the matrix of second derivatives and it is multiplying the vector $(\mathbf{r} - \mathbf{r_0})$ to get a new vector which is then dotted into $(\mathbf{r} - \mathbf{r_0})$ to get a function of **r**.

Chapter 2.7:

- Min-max for 1-variable f(s): If s_0 is a local max or min for f(s), then $f'(s_0) = 0$. Conversely, if $f'(s_0) = 0$ and $f''(s_0) > 0$, then s_0 is a local min; or if $f''(s_0) < 0$, then s_0 is a local max.
- If f(x, y) has a local min or max at $\mathbf{r}_0 = (x_0, y_0)$, then $\nabla f(\mathbf{r}_0) = \mathbf{0}$.
- Contours for local max, min and saddles.
- 2'nd derivative test: Let $\det(f'') = f_{xx}f_{yy} - f_{xy}^2$ and let $\text{tr}(f'') = f_{xx} + f_{yy}$. If $\mathbf{r}_0$ is a point where $\nabla f = \mathbf{0}$, and if $\det(f'') > 0$ and $\text{tr}(f'') > 0$ at $\mathbf{r}_0$, then $\mathbf{r}_0$ is a local minimum. If $\det(f'') > 0$ and $\text{tr}(f'') < 0$ at $\mathbf{r}_0$, then $\mathbf{r}_0$ is a local maximum. If $\det(f'') < 0$ at $\mathbf{r}_0$, then $\mathbf{r}_0$ is a saddle. If $\det(f'') = 0$ or if $\text{tr}(f'') = 0$ and $\det(f'') > 0$, then you can't tell and have to look at the function more closely. (Analogous to the failure of the 2'nd derivative test for 1-variable functions.) Note, this 2nd derivative test is equivalent to that in the book, but stated to be more symmetric between x and y.
- Note: It is not enough to check that both f_{xx} and f_{yy} are positive to insure a local minimum. For example, $f(x, y) = x^2 - 4xy + y^2$ has a saddle at the origin, but $f_{xx} = f_{yy} = 2$. (Indeed, along the line $t \to (t, t)$, $f(x = t, y = t) = t^2 - 4t^2 + t^2 = -2t^2$, so there is a local max on this line at $t = 0$.) Likewise, it is not enough to check that f_{xx} and f_{yy} are both negative to insure a local maximum. For example, $f = -x^2 + 4xy - y^2$ is a saddle with $f_{xx} = f_{yy} = -2$.
- Extreme points on the boundary: If the domain of interest has a boundary (such as the disk or a square in the plane), then you must check for extreme points on the boundary as well as those in the interior. For example, $f = x^2 + y^2$ has its minimum in the unit disk at the origin, but takes its maximum on the boundary.

Chapter 4.4 and website supplement on *Lagrange Multipliers*.

- If f(x, y) has a local min or max at $\mathbf{r}_0 = (x_0, y_0)$, then $\nabla f(\mathbf{r}_0) = \mathbf{0}$.
- 2'nd derivative test: Let $\det(f'') = f_{xx}f_{yy} - f_{xy}^2$ and let $\text{tr}(f'') = f_{xx} + f_{yy}$. If $\mathbf{r}_0$ is a point where $\nabla f = \mathbf{0}$, and if $\det(f'') > 0$ and $\text{tr}(f'') > 0$ at $\mathbf{r}_0$, then $\mathbf{r}_0$ is a local minimum. If $\det(f'') > 0$ and $\text{tr}(f'') < 0$ at $\mathbf{r}_0$, then $\mathbf{r}_0$ is a local maximum. If $\det(f'') < 0$ at $\mathbf{r}_0$, then $\mathbf{r}_0$ is a saddle. If $\det(f'') = 0$ or if $\text{tr}(f'') = 0$ and $\det(f'') > 0$, then you can't tell and have to look at the function more closely. (Analogous to the failure of the 2'nd derivative test for 1-variable functions.) Note, this 2nd derivative test is equivalent to that in the book, but stated to be more symmetric between x and y.
- Extreme points on the boundary: If the domain of interest has a boundary (such as the disk or a square in the plane), then you must check for extreme points on the boundary as well as those in the interior. For example, $f = x^2 - y^2$ has both min and max in the unit disk on the boundary.

- Method of Lagrange multipliers: If boundary is given by level set $g(x, y) = c$, then interior critical points have $\nabla f = 0$ and boundary exteme points are a subset (maybe all) of the points where $g = c$ that also obey the condition $\nabla f = \lambda \nabla g$ where λ can be any real number. Where this conditon holds, the level set of f is tangent to the constraint curve.
- λ is called the <u>Lagrange</u> <u>multiplier</u>.
- Can either find boundary extreme points by parametrizing the boundary or by Lagrange multiplier method.
- Note: Even though $\nabla f = \lambda \nabla g$ for a boundary point, that point need not be a local max or min. It could be an inflection point on the boundary. For example, if $f = x^3 - y$, the domain is the upper half plane (where $y \geq 0$), so the boundary is the x-axis. Then
  $g(x, y) = y$ and the constraint has $y = c$. Lagrange multiplier says boundary extrema are where $g(x, y) = y = 0$ and $\nabla f = (3x^2, -1) = \lambda (0, 1) = \lambda \nabla g$ where λ is a real number. This is solved by taking $\lambda = -1$ and $x = 0$ and then $y = 0$ to make $g = 0$. However, $(0, 0)$ is neither a local min nor max on the boundary. So, the lesson is that you have to still check the boundary points even though they obey the Lagrange condition, just as you must check the interior points even though $\nabla f = 0$.
- Extremal points of functions $f(x, y, z)$ of three variables. Suppose that R is a region in space with a boundary surface S, for example a ball with boundary the sphere. Suppose that f is a function on R. Problem: Find the maximum and minimum values of f on R.
- Step 1: Find all critical points inside R; points $\mathbf{r}_0$ where $\nabla f(\mathbf{r}_0) = 0$. Check which are local max or min. (There is a 3-d second derivative test, but it is for Math 21b; it involves some subtle properties of the Hessian matrix f''. In particular, it turns out that there are always three mutually orthogonal vectors $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ and associated numbers λ_1, λ_2, λ_3 such that $\mathbf{u}_j \cdot f'' = \lambda_j \mathbf{u}_j$ for each $j = 1, 2, 3$. Here, matrix multiplication is involved. The $\mathbf{u}$'s are called the 'eigenvectors' and the λ's are called the 'eigenvalues'. The critical point, $\mathbf{r}_0$ is a local min if each λ_j is positive and a local max if each is negative. Various kinds of saddles arise when the λ_j are not all of the same sign. If some are zero, then one must do an 'eyeball' test. In any event, we won't be dealing with this---Math 21b does.)
- Step 2: Look for extreme points on the boundary. If the boundary is given as a level set of some other function $g(x, y, z)$, say S is the set where $g(x, y, z) = c$ where c is a constant, then the extreme points on the boundary are always a subset of those where $g = c$ and $\nabla f = \lambda \nabla g$. Here, λ can be any real number.
- The eyeball test must be done to decide which of the points on the boundary which obey $\nabla f = \lambda \nabla g$ and which in the interior which obey $\nabla f = 0$ are the extreme points. For this purpose, just plug in the coordinates of each of these points into f and see which points give the largest and which the smallest answer.

- By the way, if the region R for the extremal problem is unbounded, then there may not be extreme points. For example, in 1-d, the function $t \to e^t$ has no extreme points on R. In 2-d, the function $(x, y) \to x + y$ has no extreme points on R^2. In 3-d, the function $(x, y, z) \to x^2yz$ has no extreme points where $x \geq 0$.
- Note that the Lagrange multiplier technique can be used to answer problems of the following sort: Find the max and min of $f(x, y, z)$ where (x, y, z) are constrained by some other function. For example, by requiring that $g(x, y, z) = c$, where g is the other function. Thus, one is looking for the max and min of f on the surface where $g = c$.

Chapter 2.8:

- The Chain Rule for composing functions of 1-variable: $f(g(t))' = f'(g(t)) g'(t)$.
- The Chain Rule for functions of two or more variables: $(x, y) \to f(x, y)$ and $t \to \mathbf{r}(t)$ where $\mathbf{r}(t) = (x(t), y(t))$ composes to give $f(x(t), y(t))$, a function of t. This is the restriction of f to the curve parametrized by t. (It is often called the 'pull-back' of f.) Then, $f(\mathbf{r}(t))' = \nabla f|_{\mathbf{r}(t)} \bullet \mathbf{r}'(t)$.
- Similarly, for $f(x, y, z)$ and $t \to \mathbf{r}(t)$, then $f(\mathbf{r}(t))' = \nabla f|_{\mathbf{r}(t)} \bullet \mathbf{r}'(t)$.
- Chain rule summary: If f is a function of x, y and z and the x, y and z are written as functions of one variable, u, or two (u, v), then $f_u = f_x x_u + f_y y_u + f_z z_u$ and $f_v = f_x x_v + f_y y_v + f_z z_v$. This is a generalization of the 1-variable rule where f is a function of just x and x just of u, so $f_u = f_x x_u$. Now, f is a function of three variables, so apply the 1-variable rule 3 times when computing f_u (pretend that v is just a constant in the definition of f. Likewise, apply the 1-variable rule 3 times when computing f_v while pretending u is constant.
- Other ways to write the composition rules: If $\mathbf{X}(u, v) = (x(u, v), y(u, v), z(u, v))$ is a map from a region in R^2 (the u-v plane) to R^3 and $f: R^3 \to R$ is a function, then $f(\mathbf{X}(u, v))$ defines a function on R^2, the u-v plane. As such, its gradient is $\nabla f(\mathbf{X})|_{(u,v)} = \nabla f|_{\mathbf{X}(u,v)} \bullet \mathbf{X}'|_{(u,v)}$ where $\mathbf{X}'$ is the matrix of derivatives of $\mathbf{X}$, namely $\begin{pmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix}$ and the composition is matrix multiplication.
- If $\mathbf{X} = \begin{pmatrix} x(u, v) \\ y(u, v) \end{pmatrix}$ maps R^2 to R^2 (the u-v plane to the x-y plane) and $\mathbf{Y} = \begin{pmatrix} u(s, t) \\ v(s, t) \end{pmatrix}$ maps R^2 to R^2 (the s-t plane to the u-v plane) then $\mathbf{X}(\mathbf{Y}(s, t))$ maps the s-t plane to the x-y plane and its 2×2 matrix of derivatives, $(\mathbf{X}(\mathbf{Y}))'|_{(s,t)}$ is $\mathbf{X}'|_{\mathbf{Y}(s,t)} \bullet \mathbf{Y}'|_{(s,t)}$ where matrix multiplication is implicit and $\mathbf{X}' = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix}$ and $\mathbf{Y}' = \begin{pmatrix} u_s & u_t \\ v_s & v_t \end{pmatrix}$.

- Generally, if **X** maps R^m to R^n and **Y** maps R^k to R^m, then these maps can be composed so as to give **X**(**Y**) whose matrix of derivatives has k rows and m columns and is $\mathbf{X}'|_{\mathbf{Y}} \cdot \mathbf{Y}'$.
- These formula are proved by looking at the linear approximations to the functions or maps involved.

Chapter 3.1

- 1-variable integration: $\int f(s)\, ds$ = limit of approximation of area under the graph by sums.
- 2-variable integration: Suppose R is a region in the x-y plane, and f is a function of x-y. Define $\iint_R f(x, y)\, dA$ as a limit of sums formed by (approximately) filling R using N tiny squares. Thus, if the squares are labeled from 1 to N, then $\iint_R f\, dA = \lim_{N \to \infty} \Sigma_{j=1,\dots,N} f(P_j)\, \Delta A_j$, where $P_j = (x_j, y_j)$ is the center of the j'th square and ΔA_j is the area of the j'th square. Here, we require that $\Delta A_j \sim$ Area(R)/N so that the squares are getting smaller as $N \to \infty$. By the way: This gives $\iint_R 1\, dA =$ Area(R).
- Triple integrals: Definition of the integral of a function f(x, y, z) over some volume V in R^3: Approximately fill V by N tiny cubes labeled from 1 to N. Then: $\iiint_V f(x, y, z)\, dV = \lim_{N \to \infty} \Sigma_{j=1,\dots,N} f(P_j)\, \Delta V_j$, where P_j is the center of the j'th cube and $\Delta V_j \sim$ Volume(V)/N is the volume of the j'th cube.
- This definition gives $\iiint_V 1\, dV =$ Volume(V).
- Interpretation of $\iint_R f\, dA$ and $\iiint_V f\, dV$ as the average value of f over either R or V.

Chapter 3.2 and website supplement on *Triple Integrals*:

- Calculation by iterated integrals: Integrals over rectangle $a \le x \le b$ and $c \le y \le d$ as $\int_a^b \left(\int_c^d f(x,y)\, dy \right) dx$ or else $\int_c^d \left(\int_a^b f(x,y)\, dx \right) dy$. That is, either do y integration first keeping x constant, or x first and y constant. The answer doesn't care which you do.
- General iterated integrals over region with boundary, say $g(x) \le y \le h(x)$ with $a \le x \le$ b. Then, the integral is equal to $\int_a^b \left(\int_{g(x)}^{h(x)} f(x,y)\, dy \right) dx$. Alternately, when $g(y) \le x \le$ h(y), and $c \le y \le d$, then $\int_c^d \left(\int_{g(y)}^{h(y)} f(x,y)\, dx \right) dy$.
- View iterated integrals as slicing the integration region by either vertical lines (and so do the y integral first) or by horizontal lines and so do the x integral first.
- The order in which you do the integrals may make a big difference in the degree of difficulty. For example, consider integrating $\sin(y^2)$ over the region where $0 \le x \le 1$ and $x \le y \le 1$.

- Iterated integrals for $\iiint_V h(x, y, z)\,dV$. In the case where V is the box where $a \le x \le b$, $c \le y \le d$, and $e \le z \le f$, then this is the integral $\int_a^b \left(\int_c^d \left(\int_e^f h(x,y,z)\,dz \right) dy \right) dx$. Here, again, the order of doing the integrals doesn't affect the answer although it may affect the degree of difficulty.

- Slicing for triple integrals: Suppose that h(x, y, z) is to be integrated over a domain where $a \le x \le b$, $c \le y \le d$ and $e(x, y) \le z \le f(x, y)$. Then, the triple integral is $\int_a^b \left(\int_c^d \left(\int_{e(x,y)}^{f(x,y)} h(x,y,z)\,dz \right) dy \right) dx$.

- In general, when you slice to do the z-integration first, you end up with an x-y integration over the 'shadow' of the region V in the x-y plane. The analog is slicing, say by vertical lines, to do $\iint_R f(x, y)\,dA$, by doing the y integration first. The result is an x-integral over the shadow of the integration region R in the x-axis. This shadow region is the region that intersects all vertical lines that pass through points of R. For example, if the function $f(x, y) = xy\ e^{-x^2-y^2}$ and R is the region of the plane where $x^2 - y^2 \ge 1$, then the shadow of R on the x-axis consists of the points on the x-axis where $x \le -1$ and also those points where $x \ge 1$. Note that in this example, the shadow region has two parts, not just one.

- Iterated integrals for $\iiint_V h(x, y, z)\,dV$. In the case where V is the box where $a \le x \le b$, $c \le y \le d$, and $e \le z \le f$, then this is the integral $\int_a^b \left(\int_c^d \left(\int_e^f h(x,y,z)\,dz \right) dy \right) dx$. Here, again, the order of doing the integrals doesn't affect the answer although it may affect the degree of difficulty.

- Slicing for triple integrals: Suppose that h(x, y, z) is to be integrated over a domain where $a \le x \le b$, $c \le y \le d$ and $e(x, y) \le z \le f(x, y)$. Then, the triple integral is $\int_a^b \left(\int_c^d \left(\int_{e(x,y)}^{f(x,y)} h(x,y,z)\,dz \right) dy \right) dx$.

- Slicing regions for triple integrals: Let V be a region in R^3, h(x, y, z) a function on V. To compute $\iiint_V h\,dV$, you can do x, y, z integrals in any order. For example, if you do x first, then for each pair (y, z), you the lower limit for the x-integral is the value, a(z, y), of x at the point where the line at constant (y, z) enters V, and the upper limit is the value, b(y, z), where the line at constant (y, z) exits V. For each pair (y, z), the 'partial' integral $\int_{a(y,z)}^{b(y,z)} h(x,y,z)\,dx$ now a number, so as (y, z) vary, you have a function which you must now integrate over the shadow of V in the y-z plane as a double integral. Here, the shadow of V is the region R in the y-z plane consisting of the points on the lines parallel to the x axis which pass through V. This integral over R should then be treated in the usual way as a double integral, resulting in $\iiint_V h\,dV = \iint_R \left(\int_{a(y,z)}^{b(y,z)} h(x,y,z)\,dx \right) dA$.

- Given a double or triple integral, know how to determine the region over which the integral is to be taken.

Chapter 3.3 and Appendix B

- Integrals in polar coordinates: How would you integrate $\iint e^{-x^2-y^2}\,dA$? Set $r = (x^2 + y^2)^{1/2}$ and $\theta = \text{Arctan}(y/x)$ as usual. Then do the integral in polar coordinates as an iterated integral. This should be easy since the integrand is independent of the coordinate θ. The only complication is that $dA = r\,dr\,d\theta$ in polar coordinates, since the area of the polar 'rectangle' where the distance from the origin ranges from r to $r + \Delta r$ and angle ranges from θ to $\theta + \Delta\theta$ is equal to $2^{-1}((r+\Delta r)^2 - r^2)\,\Delta\theta = r\,\Delta r\,\Delta\theta + 2^{-1}\Delta r^2\,\Delta\theta \sim r\,\Delta r\,\Delta\theta$ when Δr is very small.
- Thus, $\iint_R e^{-x^2-y^2}\,dA$ when $R = R^2$ in polar coordinates is the iterated integral
$$\int_0^{2\pi}\left(\int_0^{\infty} e^{-r^2}\,rdr\right)d\theta = \pi.$$
- Slicing in polar coordinates.

Chapter 3.4:

- Integrals in polar coordinates review: Remember, $dA = r\,dr\,d\theta$ where $0 \le r < \infty$ and $0 \le \theta < 2\pi$.
- Triple integrals in cylindrical coordinates: $dV = r\,dr\,d\theta\,dz$ with $x = r\cos\theta$, $y = r\sin\theta$.
- Triple integrals in spherical coordinates: $dV = \rho^2\,d\rho\sin\phi\,d\phi\,d\theta$. Here, $x = \rho\sin\phi\cos\theta$, $y = \rho\sin\phi\sin\theta$ and $z = \rho\cos\phi$, where $0 \le \rho < \infty$, $0 \le \phi \le \pi$ and $0 \le \theta \le 2\pi$. Thus, $\rho^2 = x^2 + y^2 + z^2$ is the distance from the origin, $\phi = $ latitude measured with the equator being $\pi/2$ and the north pole $\phi = 0$, and $\theta = $ longitude.
- When to do spherical versus cylindrical versus Cartesian coordinates: When the integral has an obvious symmetry.
- Slicing in cylindrical and spherical coordinates.
- When to do spherical versus cylindrical versus Cartesian coordinates: When the integral has an obvious symmetry.
- Think of switching orders of integration or using polar, cylindrical or spherical coordinates to rewrite a seemingly hard integral.

Website supplement on *Differential Equations*

- Solving the equation $\dfrac{d}{dt}p = f(p)$ for a function $p(t)$.
- There is a unique solution for each starting value p_0 for p at time $t = 0$.

- The exponential growth equation, $\frac{d}{dt}p = a\,p$ with a = constant. All solutions are of the form $p(t) = p_0\,e^{at}$ where p_0 is a constant.
- This equation arises ubiquitously due to Taylor's theorem and also because it models the growth and death of non-interacting entities.
- The tautological equation $\frac{\partial}{\partial t}u(t, x) = -\frac{\partial}{\partial x}q(t, x) + k(t, x)$ and how it arises by considering the bookkeeping of 'particles' as they are born, die and move about.
- The example where q = c u, with c = constant. This is called the advection equation; it arises when the motion of the 'particles' under consideration is due to constant motion with speed |c| of the ambient fluid.
- When k = 0 everywhere, all solutions have the form u(t, x) = f(x – c t) where f is any function of one variable.
- In the general case, the advection equation is predictive in the following sense: If you specify a function $u_0(x)$, there is a unique solution u(t, x) with $u(0, x) = u_0(x)$
- The diffusion equation: $\frac{\partial}{\partial t}u = -\frac{\partial^2}{\partial x^2}u + k(t, x)$ where μ is a positive constant. This arises for $q = -\mu\,\frac{\partial}{\partial x}u(t, x)$ in the tautological equation.
- The reason why this equation arises when the it arises when the motion of the 'particles' under consideration is random.
- A fundamental solution of the diffusion equation when k = 0 everywhere: u = $\frac{a}{t^{1/2}}e^{-x^2/4\mu t}$
- In the general case, the diffusion equation is also predictive: Give an initial function $u_0(x)$ which behaves well as $|x| \rightarrow \infty$ and there is a unique solution u(t, x) to the diffusion equation with $u(0, x) = u_0(x)$.
- The superposition principle for linear equations.
- Laplace's equation: $(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})u + k(x, y, z) = 0$ for a function u(x, y, z).
- Laplace's equation can be solved by integrating with a Green's function.

(The outline below of Chapter 5 is only relevant for the Regular and Physics sections).

Chapter 5.1:

- Vector fields: A vector valued function on R^2 or R^3: $\mathbf{F}(x, y, z) = (a(x, y, z), b(x, y, z), c(x, y, z))$. Such as the gradient of a function, ∇f. However, not all vector fields are gradients. A necessary condition for a vector field on R^2, $\mathbf{F}(x, y) = (a(x, y), b(x, y))$ is

for $a_y = b_x$ since if $a = f_x$ and $b = f_y$, then this condition holds. This turns out to be sufficient too on a region in R^2 with no holes (like the inside of the disk, as opposed to the outside. For example, $(y/(x^2 + y^2), -x/(x^2 + y^2))$ obeys $a_y = b_x$, but isn't a gradient of a function on the complement of the origin in R^2.

- Integration of vector fields along a curve: Divided by the length of the path, this measures the average of the projection of **F** along the tangent vector to the curve. Suppose γ is an unparametrized path in R^2 and R^3. Choose $N + 1$ sequential points, $\{\mathbf{r}_0, \ldots, \mathbf{r}_N\}$ along γ so that $|\mathbf{r}_i - \mathbf{r}_j|$ is small and goes to zero with $N \to \infty$. Also, make $\mathbf{r}_0$ the start and $\mathbf{r}_N$ the end of γ (choose a direction along γ for this). Then, the path integral of **F** along γ is denoted $\int_\gamma \mathbf{F} \cdot d\mathbf{r}$ and defined to be: $\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \lim_{N \to \infty} \Sigma_{1 \le j \le N}$ $\mathbf{F}(\mathbf{r}_j) \cdot (\mathbf{r}_j - \mathbf{r}_{j-1})$.

- Note that if γ is parametrized by $t \to \mathbf{r}(t)$ where t runs between 0 and T, and we break the interval [0, T] into segments of lenth $\Delta t = T/N$, we can set $\mathbf{r}_j = \mathbf{r}(j\,\Delta t)$ and thus have $\mathbf{r}_j - \mathbf{r}_{j+1} \sim \mathbf{r}'(t_j)\,\Delta t$. We then see that $\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \lim_{N \to \infty} \Sigma_{1 \le j \le N} \mathbf{F}(\mathbf{r}(t_j))\, \mathbf{r}'(t_j)\,\Delta t$ which is
$$\int_0^T \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\, dt.$$

- Hence, our path integral can be computed as a standard one variable integral if we parametrize γ. However, the answer does not depend on the parametrization as the definition surely doesn't. Thus, parametrization is a way to compute such an integral. The answer does depend on the orientation of γ --the direction chosen to go from the start point to the end point. It changes sign when you change this direction.

- Be able to compute path integrals over simple curves by parametrizing the curve in a convenient manner.

Chapter 5.2:

- Note: If $\mathbf{F} = \nabla f$ along γ for some function f, then $\int_\gamma \mathbf{F} \cdot d\mathbf{r} = f(\mathbf{P}_{end}) - f(\mathbf{P}_{start})$ where $\mathbf{P}_{start}$ is the start point and $\mathbf{P}_{end}$ is the ending point for γ. This follows using the Chain rule since $\nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = \dfrac{d}{dt} f(\mathbf{r}(t))$.

- Measure vorticity by path integrals around a closed loop: If $\mathbf{P}_{start} = \mathbf{P}_{end}$, γ is a closed loop and $\int_\gamma \nabla f \cdot d\mathbf{r} = 0$.

- Algebraic properties of the path integral: If γ_1 and γ_2 are oriented paths where γ_2 starts at the end point of γ_1, then they can be joined in the obvious way to make a new path, γ. Then, $\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_{\gamma_1} \mathbf{F} \cdot d\mathbf{r} + \int_{\gamma_2} \mathbf{F} \cdot d\mathbf{r}$. Also, if γ is traversed in the opposite direction (call the latter path γ-), then $\int_{\gamma-} \mathbf{F} \cdot d\mathbf{r} = - \int_\gamma \mathbf{F} \cdot d\mathbf{r}$. If $\mathbf{F}_1$ and $\mathbf{F}_2$ are vector fields, then for any γ and real numbers α, β, one has $\int_\gamma (\alpha\, \mathbf{F}_1 + \beta\, \mathbf{F}_2) \cdot d\mathbf{r} = \alpha \int_\gamma \mathbf{F}_1 \cdot d\mathbf{r} + \beta \int_\gamma \mathbf{F}_2 \cdot d\mathbf{r}$.

Chapter 5.3:

- Green's theorem: Suppose that γ is piecewise smooth loop with no self-intersections, and so γ is the boundary of a region, R, in the plane. Suppose that $\mathbf{F} = (a(x, y), b(x, y))$ is a vector field defined along γ and in R. Then, $\int_\gamma \mathbf{F} \bullet d\mathbf{r} = \iint_R (b_x - a_y)\, dxdy$. Here, the path integral along γ is defined using the right hand rule orientation for γ. This is the 'right hand rule' given as follows: Use your fingers on your right hand to point outward from R across γ and then the direction to move along γ is towards your thumb when your palm is flat on the plane and the thumb is at right angles to the other fingers. Thus, if γ is a circle, then the right hand rule gives the counter-clockwise direction along γ.
- If a region has holes, with outer boundary γ_0 and inner boundaries $\{\gamma_0, \ldots, \gamma_k\}$ (so there are k holes), then $\iint_R (b_x - a_y)\, dx\, dy = \sum_{0 \le j \le k} \int_{\gamma_k} \mathbf{F} \bullet d\mathbf{r}$, where the γ_j are all oriented by the 'right hand rule'. If the outer boundary is a circle and the inner boundaries also circles, the right hand rule gives the counterclockwise orientation around the outer boundary and the clockwise orientations around all of the inner boundaries. Alternately, if you use the counter clockwise orientation around all boundary circles, then the outer boundary contribution has a + sign and the inner ones all have – signs.
- Be able to translate complicated path integrals into area integrals via Green's theorem (or vice-versa) to solve seemingly complicated area or path integrals.
- Use the fact that some path integrals are independent of the chosen path to equate an integral over a complicated path with one over a simpler path. (Path independence will hold for $\mathbf{v} = (a, b)$ when $b_x - a_y = 0$.)

Chapter 5.4:

- Parametrizing surfaces: Putting on coordinates to find your way around. This constitutes a map, $\mathbf{X}$, from a portion of R^2 (say with coordinates labled (u, v)) to R^3. Thus, $\mathbf{X}(u, v) = (x(u, v), y(u, v), z(u, v))$.
- Example 1: A graph, $z = f(x, y)$ can be parametrized by $(u, v) \to \mathbf{X}(u, v) = (u, v, f(u, v))$.
- Example 2: A sphere where $x^2 + y^2 + z^2 = 1$ can be parametrized by angles (ϕ, θ) where $0 \le \phi \le \pi$ and $0 \le \theta \le 2\pi$ via $\mathbf{X}(\phi, \theta) = (\cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi)$.
- Example 3: An ellipsoid where $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ can be parametrized by the same (ϕ, θ) via the map $\mathbf{X}(\phi, \theta) = (a \cos \theta \sin \phi, b \sin \theta \sin \phi, c \cos \phi)$.
- Example 4: A donut obtained by rotating the circle $(x - 2)^2 + z^2 = 1$ around the z-axis can be parametrized by sending (φ, θ) where $0 \le \varphi \le 2\pi$ and $0 \le \theta \le 2\pi$ to $\mathbf{X}(\varphi, \theta) = ((2 + \sin \varphi) \cos \theta, (2 + \sin \varphi) \sin \theta, \cos \varphi)$.

- The tangent plane to the image of $\mathbf{X}(u, v)$ is spanned by the vectors $\mathbf{X}_u$ and $\mathbf{X}_v$. A normal vector to the surface is thus $\mathbf{X}_u \times \mathbf{X}_v$.
- Interpreting $|\mathbf{X}_u \times \mathbf{X}_v|$ as the local area expansion (or contraction) factor.

Chapter 5.5 and website supplement on *Surface Area*:

- Integration over a surface: If S is some surface in R^3, and f is a function on R^3, then $\iint_S$ f dA should give the area of the surface when f = 1, and otherwise, should give Area $\times$ (Average of f on the surface).
- Surface integrals as a limit of sums.
- If we parametrize the surface as $(u, v) \rightarrow \mathbf{X}(u, v)$ where (u, v) range over some region R in the u-v plane, then $\iint_S$ f dA = $\iint_R$ f($\mathbf{X}(u, v)$) $|\mathbf{X}_u \times \mathbf{X}_v|$ du dv, where the right hand side is our old friend, the double integral.
- Features: This is independent of parametrization.
- May have to parametrize different portions of a given surface differently. For example, the upper and lower hemispheres of a sphere if you are integrating a function over the sphere.
- Be able to set up surface integrals over simple surfaces such as spheres, cones, planes and graphs of functions.

Chapter 5.6 and website supplement on *Curl and Divergence*:

- Definition of flux of vector field through a surface: Flux of $\mathbf{F} = \iint_S \mathbf{F} \cdot \mathbf{n}$ dS, where $\mathbf{n}$ = unit length normal vector to the surface. Note that this definition requires an 'orientation' of the surface, a choice of the unit length normal. The definition changes sign if you change the normal.
- Compare with the path integral of $\mathbf{F}$ over a curve γ: $\int_\gamma \mathbf{F} \cdot d\mathbf{r}$ which can be written as $\int_\gamma \mathbf{F} \cdot \mathbf{v}$ dl, where $\mathbf{v}$ is the unit length tangent to γ in the correct direction and where dl = $|\mathbf{r}'(t)|$ dt gives the infinitesimal length.
- In the surface case, dS = $|\mathbf{X}_u \times \mathbf{X}_v|$ du dv having parametrized the surface. As $\mathbf{n} = \pm (\mathbf{X}_u \times \mathbf{X}_v)/|\mathbf{X}_u \times \mathbf{X}_v|$, this flux integral is $\iint_S \mathbf{F} \cdot \mathbf{n}$ dS = $\pm \iint \mathbf{F}(\mathbf{X}(u, v)) \cdot (\mathbf{X}_u \times \mathbf{X}_v)$ du dv given a parametrization.
- Divergence and curl: Remember the plan is to exhibit 2 and 3 dimensional versions of the fundamental theorem of calc: $\int_{a \leq t \leq b}$ f'(t) dt = f(b) – f(a) which relates integrals of derivatives to the values on the boundaries of the region. In our upcoming examples, two different 'derivatives' are used. One is the divergence of a vector field: Write $\mathbf{F}$ = (a(x, y, z), b(x, y, z), c(x, y, z)) and div($\mathbf{F}$) = $a_x + b_y + c_z$. In short hand,

- div($\mathbf{F}$) = $\mathbf{\nabla} \cdot \mathbf{F}$, where $\mathbf{\nabla}$ is short hand for the 'vector of derivatives' $\mathbf{\nabla} \equiv (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$.

- The other derivative is curl($\mathbf{F}$) = $\mathbf{\nabla} \times \mathbf{F}$ = (c_y - b_z, a_z - c_x, b_x - a_y).

- Note, curl($\mathbf{\nabla}$f) = 0 and if curl($\mathbf{F}$) = 0 on a region where all loops can be drawn in to a point without breaking, then $\mathbf{F} = \mathbf{\nabla}$f for some function f.

- Note that div(curl($\mathbf{F}$)) = 0 and if div($\mathbf{B}$) = 0, then $\mathbf{B}$ = curl($\mathbf{F}$) provided that the region has no holes. For example, $\mathbf{B} = \mathbf{\nabla}(1/\rho)$ on the complement in R^3 of the origin. Here, $\rho = (x^2 + y^2 + z^2)^{1/2}$.

- Be able to tell from a sketch vector fields which have large curl or divergence.

Chapter 5.7:

- Five integration by parts theorems:
    1. The Fundamental Theorem of Calculus: $\int_a^b$ f´(t) dt = f(b) – f(a).
    2. $\int_\gamma \mathbf{\nabla}$f•d$\mathbf{r}$ = f(P$_{end}$) – f(P$_{start}$) where γ is an oriented path and P$_{start}$ is the starting point, P$_{end}$ is the ending point.
    3. Green's theorem: If $\{\gamma_k\}_{1 \le k \le N}$ is a set of non-intersecting, closed loops in the plane that together bound a region R and if $\mathbf{F}$ = (a, b) is a vector field defined on R, then $\iint_R$ (b$_x$ – a$_y$) dA = $\sum_{1 \le k \le N} \int_{\gamma_k} \mathbf{F}$ •d$\mathbf{r}$ where the 'right hand rule' is used to pick the orientations of each γ_k.
    4. Stoke's theorem: If $\{\gamma_k\}_{1 \le k \le N}$ is a set of non-intersecting, closed loops in space that together form the boundary of a surface, S, if $\mathbf{F}$ = (a, b, c) is a vector field defined in a region containing S, then the flux integral $\iint_S$ curl($\mathbf{F}$)•$\mathbf{n}$ dS = $\sum_{1 \le k \le N} \int_{\gamma_k} \mathbf{F}$ •d$\mathbf{r}$ where $\mathbf{n}$ is a unit normal to S and the 'right hand rule' (see below) is used to pick the orientations of each γ_k.
    5. The divergence theorem: If S is a surface in space that forms the boundary of a region V, and $\mathbf{F}$ is a vector field that is defined on V, then $\iiint_V$ div($\mathbf{F}$) dV = $\iint_S \mathbf{F} \cdot \mathbf{n}$ dS, where $\mathbf{n}$ is the outward pointing, unit length normal vector to S.

- The right hand rule for Stoke's theorem: At the edge of S along γ_k, put your right hand palm down on S with the fingers pointing out. If the top of your hand is in the direction of $\mathbf{n}$, then make the thumb point at right angles to the fingers and it points in the direction of the orientation along γ_k.

- Examples of Stoke's theorem: For a surface R in the x-y plane, Stokes = Greens, the equator of a sphere and $\mathbf{F}$ = (y, -x, 0). The case $\mathbf{F} = \mathbf{\nabla}$f.

- Examples of the divergence theorem: Volume of spheres, cubes, general averages.

- Stoke's theorem for closed surfaces, $\iint_S$ curl($\mathbf{F}$)•$\mathbf{n}$ dS = 0 when S encloses a volume V as div(curl($\mathbf{F}$)) = 0.

- Using the divergence theorem (since the flux of a curl through a closed surface is zero) to equate the Stokes' theorem flux for the top half of a closed surface to minus that of the bottom half—the computation may be easier for one or the other halves.
- Be able to use the divergence theorem to turn a complicated flux integral into a volume integral and vice-versa. Be able to use it to relate the flux integral over half of a region's boundary with minus that over the other half plus the integral of the divergence over the interior. This is a handy trick that can turn a flux integral over a complicated surface into one over a much simpler surface. It is especially useful with the divergence is zero.
- Be able to use Stoke's theorem to turn flux integrals of curls over complicated surfaces into line integrals around the boundaries.

(The outline below of Rosners' Chapters 2-5 is only relevant for the BioChem sections.)

Rosner Chapter 2:

- Histograms.
- Mean, median and mode.
- Variance and standard deviation.
- Grouped data.
- Scaling data.
- Know how to compute means, medians and modes from a given data set.

Rosner Chapter 3:

- Definition of probability.
- Translate events about probability into statements about subsets and vice-versa.
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.
- $\Pr(A \cap B) = \Pr(A)\,\Pr(B)$ when A and B are independent events.
- $\Pr(\overline{A}) = 1 - \Pr(A)$.
- Conditional probabillity: Conditional probability of B given A is denoted by $\Pr(B|A)$ and is equal to $\Pr(A \cap B)/\Pr(A)$.
- Given the probabilities of two or three events, be able to calculate the probability of other related events defined in terms of union, intersection, complement and/or difference.
- Given two events and sufficient information to determine probabilities of the events and their intersection, know how to calculate the conditional probability of one event given that the other has occurred.

- Relative risk: $\Pr(B|A)/\Pr(B|\overline{A})$. Note $\Pr(B) = \Pr(B|A) \Pr(A) + \Pr(B|\overline{A}) \Pr(\overline{A})$.
- $PV^{\pm}$. If A = symptom, B = disease, then $PV^{+} = \Pr(B|A)$ and $PV^{-} = \Pr(\overline{B}|\overline{A})$.
- Sensitivity = $\Pr(A|B)$ and Specificity = $\Pr(\overline{A}|\overline{B})$.
- Bayes rule: $\Pr(B|A) = \Pr(A|B) \Pr(B)/\Pr(A)$.
- Know how to compute $PV^{\pm}$ from a given data set.
- Know how to use Bayes rule in various simple settings to relate $\Pr(B|A)$ to $\Pr(A|B)$.

Rosner Chapter 4:

- Random variables.
- Probability mass function for a discrete random variable.
- Expected value $\mu = \sum_j x_j \Pr(X = x_j)$
- Standard deviation σ where $\sigma^2 = \sum_j x_j^2 \Pr(X = x_j) - \mu^2$.
- Cumulative distribution functionss.
- Permutations and factorials.
- The binomial distribution: Probability of k successes in n independent trials of probability of success in one trial is p is $\Pr(k) = \binom{n}{k} p^k (1 - p)^{n-k}$ where $\binom{n}{k} = n!/(k! \, (n - k)!)$ and $n! = n \, (n - 1) \, (n - 2) \cdots 1$.
- Know when to use factorials and ratios of factorials to count the number of ways of picking objects at random from a given set.
- For an experiment that yields 'success' and 'failure' with specified probabilities, repeated either a specified number of times or until a 'success' or 'failure' occurs, know how to apply the binomial distribution to calculate the probability of some number of successes, the most probable number of successes or the probability that the first success (or failure) occurs before the n'th trial.
- Expected value = $n \, p$ and variance = $n \, p \, (1 - p)$ for binomial distribution
- Poisson distribution: Probability of k events in time t is $\Pr(k) = \mu^k e^{-\mu t}/k!$, where $\mu > 0$ is a constant which is determined by the circumstances. Note, that the mean and variance are both equal to μ.
- Derivation of Poisson distribution from binomial distribution.
- Know how to compute probabilities in simple situations.
- Know how and when to apply the binomial and Poisson distributions as models of statistical phenomena.

Rosner Chapter 5:

- Probability x is between numbers a and b, with a < b, is given by $\int_a^b f(x)dx$. Here, f is

the probability density function, it takes values between 0 and 1 and its integral over the whole of the real line is equal to 1.

- The cumulative distribution function, $\Pr(X < b) = \int_{-\infty}^{b} f(x)dx$.

- Given the distribution function for a continuous random variable, be able to find its density function and vice versa.

- The mean of the distribution is $\mu = \int_{-\infty}^{\infty} xf(x)dx$.

- The standard deviation, σ, is found using $\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$.

- The normal distribution with mean μ and variance σ is $\dfrac{1}{\sqrt{2\pi}\sigma} \exp\left[-\dfrac{1}{2\sigma^2}(x-\mu)^2 \right]$.

- Normal approximation of binomial and Poisson distributions. Know when such approximations are valid.

- Given a data set, and assuming that it comes from a normal distribution, be able to write down the correct one.

- Be able to use the tabulated values for the normal distribution to compute probabilities.

(The outline below if only relevant for the CS section.)

1. Sets, events and probability

- Be able to translate events about probability into statements about subsets and additive measures, and vice-versa.
- Given two subsets defined in terms of union, intersection, complement and/or difference, know how to determine whether they specify the same set, or disjoint sets; or whether one must be a subset of the other.
- Given the probabilities of two or three events, be able to calculate the probability of other related events defined in terms of union, intersection, complement and/or difference.

2. Combinatorics

- Know how to apply principles of systematic counting to determine the number of ways various events can occur. For example, be able to determine the number of distinct poker or bridge hands of a specified type, or the number of k-element subsets of an n-element set that include one or two elements of a another subset.

3. Conditional probability

- Given two events and sufficient information to determine probabilities of the events and their intersection, know how to calculate the conditional probability of one event given that the other has occurred.
- Be able to analyze the 'Monty Hall problem' and variants of it in terms of conditional probability.
- Know how to use conditional probability to analyze situations where and event A occurs, followed by an event B for which the outcomes had probabilities that depended on the outcome of A and you are required to determine the probability of A given that B occurred.

4. Binomial distributions and Bernoulli trials

- For an experiment that yields 'success' and 'failure' with specified probabilities, repeated either a specified number of times or until a 'success' or 'failure' occurs, be able to calculate the probability of some number of successes, the most probable number of successes or the probability that the first success (or failure) occurs before the n'th trial.

5. Countability and uncountability

- Be able to state the definition of countability in terms of 1-1 correspondence of sets and describe how this definition applies to a computer program that attempts to enumerate all sets in a given collection.
- Be able to use this definition to establish key results about countability, specifically:
  a) Every subset of a countably infinite set is countable.
  b) The cartesian product of two countable sets is countable.
  c) The rational numbers are countable.
  d) The collection of n-element subsets of the positive integers is countable.
  e) The collection of all subsets of the positive integers is countable.
- Using the fact that the union of a countable collection of countable sets is countable and that the Cartesian product of a finite number of countable sets is countable (which you do not need to know how to prove), be able to establish the countability of various other sets.

6. Distribution and density functions for a random variable

- Be able to explain why, for an uncountable sample space, the set of all points with positive probability must be countable.

- Given the distribution function for a continuous random variable, be able to find its density function and vice versa.
- Given a randome variable with a uniform or exponenetial distribution and a second random variable that is a specified function of the first, know how to determine the distribution function and density function for the second random variable. Also, be able to calculate the expectation of both random variables.

7. Two dimensional random variables

- Given a density function for a two dimensional random variable, be able to:
  a) Determine the value of a constant in the density function so that the total probability
     is 1.
  b) Set up and evaluate a double integral to determine the probability of an event defined
     in terms of the random variable.
  c) Set up and evaluate a double integral to determine a distribution function or expectation of a random variable that is a function of the given two dimensional random variable.

8. Simple proofs by induction

- Know how to write out inductive proofs that extend statements about 'two things' to statements about 'n-things'. For example, be able to do this for the following specific cases:
  a) If a measure is additive for the union of two disjoint sets, then it is additive for the
  u union of n disjoint sets.
  b) If sequential counting is valid for the Cartesian product of two sets, then it is valid for the Cartesian product of n sets.
  c) If a function satisfies $g(u + v) = g(u) + g(v)$, then $g(mu/n)$ can be expressed directly in terms of $g(u)$.