ENG MATHS TO REMEMBER

CALCULUS

Partial Differentiation

$$\text{Directional derivative} := f_{\hat{v}} = \nabla f \cdot \hat{v} \tag{1}$$

$$\text{Remember } \nabla f \text{ is a vector} \tag{2}$$

$$\text{Only if } f_x \text{ and } f_y \text{ are smooth does } f_{xy} = f_{yx} \tag{3}$$

$$\text{The total derivative (of u)} := du = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \tag{4}$$

where $u = f(x, y)$

$$\text{Definition of first partial derivative} := \frac{\partial f}{\partial x} = lim_{\Delta x \to \infty} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \tag{5}$$

- Multi-dimensional Taylor series. Let $\Delta x = x - x_0$ and $\Delta y = y - y_0$. Then:

$$f(x, y) = f(x_0, y_0) + f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \frac{1}{2!}[f_{xx}(x_0, y_0)\Delta x^2 + 2f_{xy}(x_0, y_0)\Delta x \Delta y + f_{yy}(x_0, y_0)\Delta y^2]...$$

ODE's

- Substitution - Use $y = \frac{x}{t}$ for ODE's that look like this $\frac{dx}{dt} = f(x/t)$
- Integration factor. You know how this works.
- For non-linear ODE's try conservation law. You can solve the ODE if, for an ODE that looks like this $q(x, t)\frac{dx}{dt} + p(x, t) = 0$, the condition $\frac{\partial q(x,t)}{\partial t} = \frac{\partial p(x,t)}{\partial x}$ holds. The function $h(x, t)$ is a conserved quantity this means that $\frac{h(x,t)}{dt} = \frac{\partial h}{\partial x}\frac{dx}{dt} + \frac{\partial h}{\partial t} = 0$. Notice this is identical to the equation above: $\frac{\partial h}{\partial x} = q(x, t)$ and $\frac{\partial h}{\partial t} = p(x, t)$. To solve the ode, integrate $q(x, t)$ with respect to x, then differentiate with respect to t to find the constants. Finally set $h(x, t) = E$ or any constant and solve for x. Make sure your left with only one constant.

Classification of ODE's:
Highest derivative gives order. If the derivatives of dependent variable or the dependent variable itself is non-linear then the ODE is non-linear. Move all dependent terms to left hand side of the equation, if it is equal to zero its homogoneous, otherwise its non-homogenous.

The three types of solutions to higher order differential equations with constant coefficients: (Ansatz is $x(t) = Ae^{mt}$)

1) characteristic polynomial has two real solutions (roots).

$$m = a, b \tag{6}$$
$$x_C(t) = Ae^{at} + Be^{bt} \tag{7}$$

2) characteristic polynomial has only one solution

$$m = a \tag{8}$$
$$x_C(t) = Ae^{at} + B\underline{t}e^{at} \tag{9}$$

3) characteristic polynomial has two complex solutions

$$m = a \pm bj \tag{10}$$
$$x_C(t) = e^{at}(Acos(bt) + Bsin(bt)) \tag{11}$$

For particular x_P solutions just use your brain and think of an ansatz. The final solution is $x(t) = x_P(t) + x_C(t)$, i.e. the sum of the complementary function and particular integral.

Resonance - I think it has to do with the particular solution having a similar frequency to that of the complementary function. Similarly if you have an output to an ODE $f(t)$ that has the same frequency as the complementary function, use $A\underline{t}cos(\omega t) + B\underline{t}sin(\omega t)$ as your ansatz.

Systems of ODE's. (State-space form). Any system of linear homogeneous differential equations can be written in the form $\frac{dx}{dt} = \mathbf{A}\underline{x}$. $\underline{x} = \sum \underline{v}e^{\lambda t}$ is the general solution, where $\underline{v}$ are the eigenvectors and λ is the corresponding eigenvalue. They are used for large calculations that can only be solved numerically.

PROBABILITY

$$\text{Mutually exclusive} := P(A \cap B) = 0 \tag{12}$$
$$\text{Independent} := P(A \cap B) = P(A)P(B) \tag{13}$$
$$\text{Conditional Probability} := P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{14}$$

- Try using contingency tables.
- If B and C are mutually exclusive $P(B \cap C) = 0$ and $P(B \cap A) \neq 0$ and $P(C \cap A) \neq 0$ then $P((B \cup C)|A) = P(B|A) + P(C|A)$
- Some properties of $E(X)$ are $E(aX) = aE(X)$, $E(X + b) = E(X) + b$ and $E(X + Y) = E(X) + E(Y)$. Use these if you want to show that $\sigma^2 = E(X^2) - \mu^2$

Discrete Random Variables

$$Probability \quad function := P_X(x) = P(X = x)$$
$$Mean := \mu = E(X) = \sum x P_X(x)$$
$$Cumulative \quad distribution \quad function := F_X(x) = P(X \leq x)$$
$$Median := m \text{ is such that } P(X \leq m) \geq \frac{1}{2} \text{ and } P(X \geq m) \geq \frac{1}{2}$$
$$Variance := \sigma^2 = E(X^2) - \mu^2 = \sum x^2 P_X(x) - \mu^2$$

Continuous Random Variables

$$Probability \quad density \quad function \quad (pdf) := f_X(x)$$
$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f_X(x)dx$$
$$Cumulative \quad distribution \quad function := F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$
$$Mean := E(X) = \int x f_X(x)dx$$
$$Median := \int_0^m f_X(x)dx = \frac{1}{2}$$
$$Mode := \frac{df_X(x)}{dx} = 0$$
$$Variance := \sigma^2 = E(X^2) - E(X)^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx - \mu^2$$

Distributions

Discrete

- Bernoulli distribution - for two outcomes $P_X(0) = 1 - p$ and $P(1) = p$.
- Geometric distribution - for finding the first successful occurrence. X is the number of trials (n). $P(X = n) = (1-p)^{(n-1)}p$. $F_X(n) = 1-(1-p)^n$. Mean = p and Variance = $p(1-p)$
- Binomial distribution - for multiple successes (k) and failures. X is the number of successes (k).n is the number of trials. $P_X(k) = \frac{n!}{k!(n-k)!}$. Mean = np. Variance = $np(1-p)$. Just like Bernoulli except multiplied by n.
- Poisson distribution - for large n and small p Bernoulli distribution. It's basically an approximation. $P_X(k) = \frac{\lambda^k}{k!}e^{-\lambda}$, where $\lambda = np$ or the mean *rate* per unit time. Mean = λ. Variance = λ.

Continuous

- Exponential distribution - For studying the time between successive events. Let X be the time between two occurrences. Then pdf $f_X(t) = \lambda e^{-\lambda t}$ and cdf $F_X(t) = 1 - e^{-\lambda t}$, for $t > 0$. Mean $= \frac{1}{\lambda}$. Variance is $\frac{1}{\lambda^2}$.You can use Poisson as well.

- Normal Distribution $X \sim N(\mu, \sigma^2)$ - for a lot of things. $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Converting to standard Normal Distribution. $Z \sim N(0,1)$ and $Z = \frac{X-\mu}{\sigma}$. So if $X \sim N(4,4)$ for example, then $P(X \leq 6.7) = P(Z = \frac{X-\mu}{\sigma} \leq \frac{6.7-4}{2})$. Find the Z values in the provided table.
- Central limit theorem - approximate a binomial with a normal distribution. Then $X \sim N(\mu, \sigma^2)$ becomes $X \sim N(np, np(1-p))$

COMPLEX NUMBERS

$$\text{De Moivre's formula} := (cos(\theta) + jsin(\theta))^n = cos(n\theta) + jsin(n\theta) \tag{15}$$

Use when asked to convert $sin(n\theta)$ or $cos(n\theta)$ to powers of cos and sin. E.g. $cos(5\theta) = Re(cos(5\theta) + jsin(5\theta)) = Re((cos(\theta) + jsin(\theta))^5)$ (Then use binomial expansion).
To convert powers of sin and cos in terms of $sin(k\theta)$ and $cos(k\theta)$, use the fact that $e^{j\theta} = cos(\theta) + jsin(\theta)$ and $e^{-j\theta} = cos(\theta) - jsin(\theta)$
Remember that $cos(jx) = cosh(x)$ and $sin(jx) = jsinh(x)$. If you get lost just set $\theta = jx$ and then use the above equations for $e^{j\theta}$. Use it for things like $cos(z) = 2$

VECTORS

Projection of one vector onto another

$$|\overrightarrow{OP}| = \frac{\underline{a} \cdot b}{|b|} = \underline{a} \cdot \hat{b} \tag{16}$$

Cross product

$$a \times b = |a||b|sin(\theta)\underline{\hat{n}} = det \begin{bmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \tag{17}$$

$$\frac{(a \times b)}{2} = \text{Area of a triangle} \tag{18}$$

$$a \times (b + c) = (a \times b) + (a \times c) \tag{19}$$

$$(ka) \times (pb) = kp(a \times b) \tag{20}$$

Vector equation of a sphere

$$|r - a| = radius \tag{21}$$

$$a = \text{position of centre of sphere} \tag{22}$$

$$r = \text{position vector} \tag{23}$$

Scalar Triple Product (Magnitude gives volume of parallelepiped)

$$a \cdot (b \times c) = det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \tag{24}$$

$$\tag{25}$$

MATRICES

Useful things to know

$$(n \times m) \times (p \times q) \text{ only works if } p = n \text{ and gives a matrix of size } m \times q$$
$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$
$$\text{Area/Volume of transformed shape by matrix} = \text{Original area/volume} \times \text{determinant of matrix}$$
$$det(A) = \sum a_{i,j} A_{i,j}$$
$$\textbf{Cofactor} A_{i,j} = (-1)^{i+j} M_{i,j} \textbf{(Minor)}$$
$$det(\mathbf{A}^T) = det(\mathbf{A})$$
$$det(\mathbf{AB}) = det(\mathbf{A}) det(\mathbf{B})$$
$$adj(\mathbf{A}) = \text{The transpose of the matrix of cofactors}$$
$$tr(\mathbf{A}) = \sum a_{i,i} = \sum \lambda$$
$$det(\mathbf{A}) = \prod \lambda$$

Eigenvalues of A^{-1} are $\frac{1}{\lambda}$. For kA they are $k\lambda$ and for A^k they are λ^k. For A^T they are the same as for A, namely just λ.
If Eigenvalues are complex, then matrix rotates as well as stretching. To analyze, convert λ into exponential form $re^{j\theta}$ and r gives stretch factor and θ gives rotation. https://www.youtube.com/watch?v=yj1eYJyOFac
If $det(\mathbf{A}) \neq 0$ then A^{-1} exists and there is a unique solution solution to the system of equations $\mathbf{A}\underline{x} = \underline{b}$ namely $\underline{x} = A^{-1}\underline{b}$
Cayley-Hamilton Theorem - Matrices satisfy their own characteristic equations.

- Unique solution has full rank and $det(A) \neq 0$
- There is no solution when $Rank(A) \neq Rank(A|b)$
- There is a family of solutions if $Rank(A) = Rank(A|b)$ (under-determined)
- $(Rank(A) = Numberofrows - Null(A))$

Use the following property to find inverses: $AA^{-1} = \mathbf{I}$ and then use row elimination.
Matrices are distributive and associative, but only sometimes commutative.

FUNCTIONS

- X is a set called the domain
- Y is a set called the co-domain
- $f : X \to Y$ is a function, if it maps a unique value from set X onto set Y.
- Injective (one-to-one) means every value in set Y is mapped to by one point in set X. It also means some of the range (co-domain) is not covered.
- Surjective(onto) means every value in set Y is mapped to by at least one point in set X. Range is completely covered.
- Bijective means it is both one-to-one and onto. If $f(x)$ is bijective then it is invertible and its inverse is bijective too.
- Even $:= f(x) = f(-x)$
- Odd $:= f(-x) = -f(x)$
- Periodic $:= f(x + T) = f(x)$

Intermediate value theorem. Basically says that if a function is continuous on an interval $[a, b]$ then $f(x)$ takes every value between $f(a)$ and $f(b)$. No shit sherlock.

A function that is smooth is continuous and differentiable throughout its domain

$$\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = f(a) \tag{26}$$

$$\lim_{x \to a^+} f'(x) = \lim x \to a^- f'(x) = f'(a) \tag{27}$$

BASIC CALCULUS

- positive curvature := convex
- negative curvature := concave
- integration "smooths out" small errors, i.e. it can turn a discontinuous function into something continuous.

The Riemann Integral (Does not work for improper integrals)

$$\int_a^b f(x)dx = \lim_{N \to \infty} \frac{(b-a)}{N} \sum_{n=1}^{N} f(a + \frac{n(b-a)}{N}) \tag{28}$$

$$\Delta x = \frac{b-a}{N} \tag{29}$$

Partial Fractions. Always remember to have as many fractions as factors in the denominator. E.g.

$$\frac{1}{(x-2)^2(x-1)} = \frac{A_1}{(x-2)} + \frac{A_2}{(x-2)^2} + \frac{A_3}{(x-1)} \tag{30}$$

NUMERICAL METHODS

- Euler Method.(Explicit) Let $\frac{dx}{dt} = f(x,t)$ be a non-linear ODE. Then the Euler method is as follows: $x_{n+1} = x_n + h\frac{dx}{dt} = x_n + hf(x_n, t_n)$. h is the step-size in the independent variables (in this case t) direction. You can also use it for higher-order ODE's but you have to convert it into state-space form.

- Backwards Euler (implicit) $x_{n+1} = x_n + hf(x_{n+1}, t_{n+1})$. $O(h)$.

- Errors. Useful to know how accurate the numerical solutions are. There are two kind of errors. Local error $= x_1 - x(t_1)$, i.e. the error after one step, and the global error $= x_n - x(t_n)$, the error after n steps. Normally if the local error gets smaller (by decreasing the step-size) the global error gets smaller. Not the case for stiff-problems though.

- For stiff ode's use an implicit method.

- step-size. You can test for which step-size various numerical methods will break by solving for x_{n+1} with an ODE that you know the solution to.

- Big O notation. A classification of how fast the error decays, i.e. how accurate the method is. For the Euler Method the error decays at a rate $O(h)$, which means it decays linearly with decreasing step-size. For the explicit midpoint method, $O(h^2)$, having the step-size will quarter the error. Runge-Kutta Method is $O(h^4)$ accurate.

- Fixed Point iteration. For solving roots of transcendental equations (anything besides polynomials) numerically, because they cannot be solved analytically. Use $x_{n+1} = g(x_n)$ and iterate until it converges.
  - A necessary condition for convergence to a root L is $|g'(L)| < 1$.
  - If $0 < g'(L) < 1$ convergence is monotonic
  - If $-1 < g'(L) < 0$ convergence is oscillatory
  - If a root diverges, you can sometimes find a rearranged iteration scheme that find a converged root.
  - Error. $E_{n+1} \approx g'(L)E_n$, where $E_n = x_n - L$.
  - rate of converges $r = |g'(L)|$

- Intermediate value theorem (IMVT) can be used to show a root exists. Its a sufficient condition not a necessary condition.

- Bisection method uses this principle to determine a root by constantly having the distance between the two opposite points $f(a)$ and $f(b)$.
- Converges linearly with rate 1/2.

- Newton-Raphson Method $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. You have to set $f(x) = 0$ and then you can find $f'(x)$
  - May jump outside interval before converging
  - May diverge
  - Converges quadratically

Series

- Series worth remembering: $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n} = 1$, $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ diverges,
- If the sequence of partial sums of a series converges, then the series converges too.
- Look at formula sheet for sum of geometric series.
- Tests
  - **Integral test**. $\int_N^\infty f(x)dx \leq \displaystyle\sum_{n=N}^{\infty} f(n) \leq \int_N^\infty f(x)dx + f(N)$
  - $\displaystyle\sum_{n=N}^{\infty} a_n$ diverges if $\lim_{n\to\infty} \neq 0$
  - **Alternating series** $\displaystyle\sum_{n=N}^{\infty} a_n$ converges if:
    * a_n is of alternating sign
    * $\lim_{n\to\infty} |a_n| = 0$
    * $|a_{n+1}| < |a_n| \forall n$
  - **Comparison test**. $\displaystyle\sum_{n=N}^{\infty} a_n$ converges if:
    * $b_n \geq a_n \geq 0$ and $\displaystyle\sum_{n=N}^{\infty} b_n$ converges
    * If $a_n \geq b_n \geq 0$ and $\displaystyle\sum_{n=N}^{\infty} b_n$ diverges, then $\displaystyle\sum_{n=N}^{\infty} a_n$ diverges
  - **Ratio test**
    * If $\lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$, then $\displaystyle\sum_{n=N}^{\infty} a_n$ is absolutely convergent

* If $\lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| > 1$, then $\displaystyle\sum_{n=N}^{\infty} a_n$ is divergent
* If $\lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| = 1$, then the test gives no information.

Sequences

- Definition. A sequence a_n converges to a limit L if $\forall\epsilon > 0$ there exists an N such that $|a_n - L| < \epsilon$, $(L - \epsilon < a_n < L + \epsilon)$ $\forall n > N$. ϵ is the distance from the limit L. So for example if we want to get to $\frac{1}{100}th$ of the limit of $1/n$, we need $\frac{1}{100} < \frac{1}{n} < \frac{1}{100}$ since $L = 0$, which means that for any $n > N = 100$, we we'll be within $\frac{1}{100}th$ of the limit.

- Recursive sequence - Each term is defined by previous term $a_{n+1} = g(a_n)$. If the sequence converges, for large n this means $a_n = g(a_n)$ so limit $L = g(L)$. You can use this to find the limit of a recursive sequence by setting a_{n+1} and a_n to L. Great for fixed point iteration.

Radius of Convergence. Simply use the ratio test on a power series and find the values for x in which the series is in between -1 and 1. E.g. You might get $|x| < 1$ from the ratio test. This means the radius of convergence is $-1 < x < 1$. To find the interval of convergence plug in the upper and lower bounds for x and see whether they converge or diverge.