Proofs of identities in FM

On The Floor

9 May 2016

1 Binomial expansion (P3)

$$(1+x)^k = 1 + \sum_{n=1}^{\infty} \left(\frac{\prod_{i=0}^{n-1}(k-i)}{n!} x^n \right) \text{ for } k \in \mathbb{Q}.$$

Proof: Notice, at first, that for nonnegative integers k and n, $\binom{k}{n} = \frac{k!}{n!(k-n)!} = \frac{k(k-1)\cdots(k-n+1)}{n!}$.

For $n \geq 1$ this is also equivalent to $\left(\frac{\prod_{i=0}^{n-1}(k-i)}{n!} \right)$. We temporarily extend this notation of $\binom{k}{n}$ to all real numbers k and all nonnegative integers n, so what we need to prove is $(1+x)^k = \sum_{n=0}^{\infty} \binom{k}{n} x^n$.

We split the problem into few steps, each expanding the set of the numbers for which the above statement is true.

1. $k \in \mathbb{N}_0$.
   Assume $k \geq 1$ since $k = 0$ is trivial. Write the expression as the product of k identical factors $1+x$, and we know that for $i \in [0, k]$ we have $\binom{k}{i}$ ways to choose x for i times and 1 for $k-i$ times. This yields the coefficient of x^i as $\binom{k}{i}$. For x^i where $i > k$ simply note that the term $\frac{\prod_{j=0}^{i-1}(k-j)}{i!}$ contains factor 0 when $j = k$, so we are done for this case.

2. $k \in \mathbb{Z}$.
   Proof: Let's do induction, each time reducing k. For base case $k = -1$ notice that $\frac{1}{1+x}$ is equal to the sum to infinity $1 - x + x^2 - x^3 + \cdots = \sum_{n=1}^{\infty}(-1)^n x^n = \sum_{n=1}^{\infty} \frac{(-1)(-2)\cdot\cdots\cdot(-n)}{(1)(2)\cdot\cdots\cdot(n)} x^n$
   $= \sum_{n=0}^{\infty} \binom{-1}{n} x^n.$

Inductive step: Suppose that $(1+x)^k = \sum_{n=0}^{\infty} \binom{k}{n} x^n$ for some $k < 0$. Differentiating both sides yields $k(1+x)^{k-1} = \sum_{n=0}^{\infty} \binom{k}{n} n x^{n-1}$, or $(1+x)^{k-1} = \sum_{n=0}^{\infty} \frac{n}{k} \binom{k}{n} x^{n-1} = \sum_{n=1}^{\infty} \frac{n}{k} \cdot \frac{k(k-1)\cdots(n-k+1)}{1\cdot 2\cdot\cdots\cdot n} x^{n-1} = \sum_{n=1}^{\infty} \frac{(k-1)\cdots(n-k+1)}{1\cdot 2\cdot\cdots\cdot(n-1)} x^{n-1} = \sum_{n=1}^{\infty} \binom{k-1}{n-1} x^{n-1}$
$= \sum_{n=0}^{\infty} \binom{k-1}{n} x^n$, as desired.

3. $k \in \mathbb{Q}$.

Proof: First, we prove that $\left(\sum_{n=0}^{\infty} \binom{k}{n} x^n\right)\left(\sum_{n=0}^{\infty} \binom{l}{n} x^n\right) = \left(\sum_{n=0}^{\infty} \binom{k+l}{n} x^n\right)$. Notice that this identity already holds for $k, l \in \mathbb{Z}$, since from the frst two subproblems we have $\left(\sum_{n=0}^{\infty} \binom{k}{n} x^n\right)\left(\sum_{n=0}^{\infty} \binom{l}{n} x^n\right) = (1+x)^k \cdot (1+x)^l$, while $\left(\sum_{n=0}^{\infty} \binom{k+l}{n} x^n\right) = (1+x)^{k+l}$.
To prove this, we need to verify that for each $n \geq 0$, coefficient of x^n is the same for LHS and RHS. For RHS it is $\binom{k+l}{n}$. For LHS this is $\sum_{i=0}^{n} \binom{k}{i}\binom{l}{n-i}$. Observe also that $\binom{k+l}{n} - \sum_{i=0}^{n} \binom{k}{i}\binom{l}{n-i}$ is a polynomial in variables k and l if n is fixed, so name it $P(k,l)$. Our aim is to prove that $P(k,l)$ is identically zero.
The fact $P(k,l) \equiv 0$ for $k, l \in \mathbb{Z}$ already follows from above. Now fix k as arbitrary integer and vary l, we know that $\binom{k+l}{n} - \sum_{i=0}^{n} \binom{k}{i}\binom{l}{n-i}$ is a polynomial in variable l and has degree at most n. However, every integer is the root of this monovariable polynomial, it has infinitely many roots and hance must be identically zero. Also that for an $k, l \in \mathbb{R}$, $P(k,l) = P(l,k)$ (P is a symmetric polynomial) so we can assert that $P(k,l) \equiv 0$ whenever one of k, l is an integer. Now turn back to $P(k,l)$ again, and fix k as any real number and vary l (so $P(k,l)$ is again a polynomial with variable l), and this time $P(k,l) = 0$ for any integer l (again!) Therefore this monovariable polynomial must be identically zero, again, and we have $P(k,l) = 0$ for any pair of $k, l \in \mathbb{R}$.

Finally, let $k = \frac{p}{q}$, and $\left(\sum_{n=0}^{\infty} \binom{k}{n} x^n\right)\left(\sum_{n=0}^{\infty} \binom{l}{n} x^n\right) = \left(\sum_{n=0}^{\infty} \binom{k+l}{n} x^n\right)$ means that $\left(\sum_{n=0}^{\infty} \binom{k}{n} x^n\right)^q = \left(\sum_{n=0}^{\infty} \binom{kq}{n} x^n\right) = \left(\sum_{n=0}^{\infty} \binom{p}{n} x^n\right) = (1+x)^p$ (since p is an integer and we already prove this grand identity true for all integers). This means $\left(\sum_{n=0}^{\infty} \binom{k}{n} x^n\right) = (1+x)^{\frac{p}{q}} = (1+x)^k$, or possibly $-(1+x)^k$ if q is even. Nevertheless, comparing the constant term (LHS=RHS=1) yields that $-(1+x)^k$ is impossible, so we are done.

Comment 1: It is possibe to establish $P(k,l) \equiv 0$ by algebraic expansion, but too messy.
Comment 2: We can jump from problem 1 directly to problem 3, but the proof of problem 2 shows the beauty of differentiation.

2 Calculus.

1. If a polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ has roots $z_1, z_2, \cdots, z_k$, each with multiplicity $b_1, b_2, \cdots, b_k$ respectively (so $b_1 + b_2 + \cdots + b_k = n$ and $b_i \geq 1$), then the equation $a_n \frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1 \frac{dy}{dx} + a_0 y = 0$ has general solution $y = \sum A_{ij}x^{j-1}e^{z_i x}$ for $i = 1, 2, \cdots, k$ and $j = 1, 2, \cdots b_i$ and for A_{ij} arbitrary constant.
Proof: Name the polynomial P. First, notice that this general solution has exactly n parameters (or n degree of freedom) because we can decide arbitrarily on what should the values $f(0), f'(0), f''(0), \cdots f^{(n-1)}(0)$ and let the rest follow according to our differential equation (we can't go further than this since the differential equation also implies $a_n \frac{d^{n+k}y}{dx^{n+k}} + a_{n-1}\frac{d^{n-1+k}y}{dx^{n-1+k}} + \cdots + a_1 \frac{d^{k+1}y}{dx^{k+1}} + a_0 \frac{d^k y}{dx^k} = 0$ for any $k \geq 0$) and these n degrees of freedom gives n parameters, according to Maclaurin's theorem. Hence it suffices to show that $x^{j-1}e^{z_i x}$ works (we can omit any constant since that will only multiply LHS by A_{ij}.

   Now let's investigate what happens as we differentiate $x^k e^{ax}$ in general. We claim, by induction, that $\frac{d^n}{dx^n}(x^k e^{ax}) = a^n x^k e^{ax} + \sum_{i=1}^{n} a^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1)x^{k-i}e^{ax}$ (notice that when $i > k$ such terms contain factor $k-(k+1)+1 = 0$ so it can be disregarded (in other words, no term has factor $x^j e^{ax}$ for $j < 0$).

   Base case $n = 0$ is trivial. Now assume that $\frac{d^n}{dx^n}(x^k e^{ax}) = a^n x^k e^{ax} + \sum_{i=1}^{n} a^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1)x^{k-i}e^{ax}$, then $\frac{d^{n+1}}{dx^{n+1}}(x^k e^{ax}) = \frac{d}{dx}\left(a^n x^k e^{ax} + \sum_{i=1}^{n} a^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1)x^{k-i}e^{ax}\right) = a(a^n x^k e^{ax}) + k a^n x^{k-1}e^{ax} + \sum_{i=1}^{n} a(a^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1)x^{k-i}e^{ax}) + \sum_{i=1}^{n}(k-i)(a^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1)x^{k-i-1}e^{ax})$. Now the term $x^k e^{ax}$ has coefficient a^{n+1} and $x^{k-i}e^{ax}$ has coefficient $a^{n-i+1}\binom{n}{i}k(k-1)\cdots(k-i+1) + a^{n-i+1}k(k-1)\cdots(k-i+1)\binom{n}{i-1} = \binom{n+1}{i}a^{n-i+1}k(k-1)\cdots(k-i+1)$ so $\frac{d^{n+1}}{dx^{n+1}}(x^k e^{ax}) = a^{n+1}x^k e^{ax} + \sum_{i=1}^{n+1} a^{n+1-i}\binom{n+1}{i}k(k-1)\cdots(k-i+1)x^{k-i}e^{ax}$, consistent with our hypothesis.

   To complete our proof, we need to verify that the coefficient of $x^{k-i}e^{ax}$ is 0 is the differential equation (for $i \leq k$). Now the coefficient of $x^{k-i}e^{ax}$ in $\frac{d^n}{dx^n}(x^k e^{ax})$ is $a^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1)x^{k-i}e^{ax}$, so for $a = z_j$ (a root of the original polynomial) with $k \leq b_j - 1$ we need $a_n z_j^{n-i}\binom{n}{i}k(k-1)\cdots(k-i+1) + a_{n-1}z_j^{n-1-i}\binom{n-1}{i}k(k-1)\cdots(k-i+1) + \cdots + a_1 z_j^{1-i}\binom{1}{i}k(k-1)\cdots(k-i+1) + a_0 z_j^{-i}\binom{0}{i}k(k-1)\cdots(k-i+1) = 0$. (Of course $\binom{k}{l} = 0$ for $k < l$). Observe that we actually assumed that $z_j \neq 0$; this limit case can be established easily wherby the general solution contains term x^{ϵ} for some ϵ, (so the polynomial is divisible by $x^{\epsilon+1}$ and $a_i = 0$ for $i \leq \epsilon$. But $\dfrac{d^i}{dx^i}x^\epsilon = 0$ for $i > \epsilon$ so $a_i\dfrac{d^i}{dx^i}x^\epsilon = 0$ for any i. Now with this assumption the above is same as proving that $\displaystyle\sum_{p=0}^{n} a_p z_j^{p-i}\binom{p}{i}$=0.

To prove this, observe that the root x_j has multiplicity of more than k times, and $i \leq k$. This means that $(x - z_j)^{k+1}$ divides the polynomial P, and it is well-known that $\dfrac{d^\alpha}{dx^\alpha}P(x)$ is divisible by $(x - z_j)^{k+1-\alpha}$ for $\alpha \leq k$ so substituting i into alpha yields that $\dfrac{d^\alpha}{dx^\alpha}P(x)$ is divisible by $(x - z_j)^{k+1-i}$ (and $k + 1 - i \geq 1$). So $\dfrac{d^i}{dx^i}P(x) = 0$ when $x = z_j$. Also observe that the t-th derivative of x^u is $u(u-1)\cdots(u-t+1)x^{u-t}$, so $0 = \dfrac{d^i}{dx^i}P(x) = \dfrac{d^i}{dx^i}\displaystyle\sum_{p=0}^{n} a_p x^p$=$\displaystyle\sum_{p=0}^{n} p(p-1)\cdots(p-i+1)a_p x^{p-i}$=$\displaystyle\sum_{p=0}^{n} i!p(p-1)\cdots(p-i+1)a_p x^{p-i}$=$i!\displaystyle\sum_{p=0}^{n}\binom{p}{i}a_p x^{p-i}$ (for $x = z_j$), as claimed.

3 Matrices

1. Row space and column space of any matrix have the same dimension.

   Proof: Let dimension of column space be n. W.L.O.G. let $\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix}$, $\begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{pmatrix}$,$\ldots$, $\begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix}$ be pairwise linearly independent. Then for $j > n$, $a_{ij} = \sum_{k=1}^{n} \lambda_{kj} a_{ik}$ for some real numbers $\lambda_{kj}a_{ik}$. Multiplying each element in row i by $\frac{1}{a_{i1}}$ we have:

$$\begin{pmatrix} 1 & \dfrac{a_{12}}{a_{11}} & \cdots & \dfrac{a_{1n}}{a_{11}} & \{\displaystyle\sum_{k=1}^{n}\lambda_{kj}\frac{a_{1k}}{a_{11}}\} \\ 1 & \dfrac{a_{22}}{a_{21}} & \cdots & \dfrac{a_{2n}}{a_{21}} & \{\displaystyle\sum_{k=1}^{n}\lambda_{kj}\frac{a_{2k}}{a_{21}}\} \\ & & \vdots & & \\ 1 & \dfrac{a_{m2}}{a_{m1}} & \cdots & \dfrac{a_{mn}}{a_{m1}} & \{\displaystyle\sum_{k=1}^{n}\lambda_{kj}\frac{a_{mk}}{a_{m1}}\} \end{pmatrix}$$

   where $\{\}$ means that j runs from $n+1, n+2, \ldots, g$ where g is the number of columns of this matrix. For convinience we name j as anything between $n+1$ and g, inclusive. (We assume that $a_{i1} \neq 0$. If it is, then just swap the whole row with some 'lower' rows with nonzero first entry.)

Now r.e.f. for the first element in each row yields

$$\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} & \{\sum_{k=1}^{n} \lambda_{kj} \frac{a_{1k}}{a_{11}}\} \\ 0 & \frac{a_{22}}{a_{21}} - \frac{a_{12}}{a_{11}} & \dots & \frac{a_{2n}}{a_{21}} - \frac{a_{1n}}{a_{11}} & \{\sum_{k=1}^{n} \lambda_{kj} (\frac{a_{2k}}{a_{21}} - \frac{a_{1k}}{a_{11}})\} \\ & & \vdots & & \\ 0 & \frac{a_{m2}}{a_{m1}} - \frac{a_{12}}{a_{11}} & \dots & \frac{a_{mn}}{a_{m1}} - \frac{a_{1n}}{a_{11}} & \{\sum_{k=1}^{n} \lambda_{kj} (\frac{a_{mk}}{a_{m1}} - \frac{a_{1k}}{a_{11}})\} \end{pmatrix}$$

If $a'_{xy} = \frac{a_{xy}}{a_{x1}} - \frac{a_{1y}}{a_{11}}$ then the original matrix becomes

$$\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} & \{\sum_{k=1}^{n} \lambda_{kj} \frac{a_{1k}}{a_{11}}\} \\ 0 & a'_{22} & \dots & a'_{2n} & \{\sum_{k=1}^{n} \lambda_{kj} a'_{2k}\} \\ & & \vdots & & \\ 0 & a'_{m2} & \dots & a'_{mn} & \{\sum_{k=1}^{n} \lambda_{kj} a'_{mk}\} \end{pmatrix}$$

Replacing each a'_{xy} with a_{xy} yields we can assume that the first element of eaach (except the first) row is zero.

Repeating the r.e.f. process yields:

$$\begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1(n-1)}}{a_{11}} & \frac{a_{1n}}{a_{11}} & \{\sum_{k=1}^{n} \lambda_{kj} \frac{a_{1k}}{a_{11}}\} \\ 0 & a_{22} & \dots & a_{2(n-1)} & a_{2n} & \{\sum_{k=1}^{n} \lambda_{kj} a_{2k}\} \\ & & \vdots & & & \\ 0 & 0 & \dots & 0 & a_{nn} & \{\sum_{k=1}^{n} \lambda_{kj} a_{nk}\} \\ & & \vdots & & & \end{pmatrix}$$

But notice that the first n elements of j-th row ($j > n$) are all zero. It follows that $\sum_{k=1}^{n} \lambda_{kj} a_{ik}$ ($\forall 1 \leq i \leq i$) must be zero. So all other elements must be zero and we have columns onsisting entirely of zeroes after the n-th row.

This yields that, the dimension of row space cannot exceed n (some of the rows 1 to n might be entirely zero, in which case we know that dimension of row space $< n$), i.e. cannot exceed the dimension of column space. Similarly the dimension of column space cannot exceed the dimension of row space. Hence they are equal.

2. Multplicative associativity of matrices: $(AB)C = A(BC)$.
   Proof: Let's check the dimension of each matrix for the equation to be valid. Let the

dimensions of A, B, C be $a_1 \times a_2$, $b_1 \times b_2$ and $c_1 \times c_2$, respectively. For AB to be valid we must have $a_2 = b_1$ and the resulting matrix has dimension $a_1 \times b_2$. For $(AB)C$ to be valid we need $b_2 = c_1$ and the matrix at LHS has dimension $a_1 \times c_2$. A similar check at RHS yields $a_2 = b_1, b_2 = c_1$ and resulting matrix with dimension $a_1 \times c_2$, hence the dimension check is done.

Now let's investigate every single entry x at i-th row and j-th column. Denote $[ab]_{xy}$ as the element of i-th row and j-th column of matrix AB; define $[bc]_{xy}$ similarly. Now for $(AB)C$ this entry is $\sum_{k=1}^{b_2=c_1} [ab]_{ik} c_{kj} = \sum_{k=1}^{b_2=c_1} (\sum_{d=1}^{a_2=b_1} a_{id} b_{dk}) c_{kj} = \sum_{1 \leq d \leq a_2, 1 \leq k \leq b_2} a_{id} b_{dk} c_{kj}$.
For $A(BC)$ this entry is $\sum_{d=1}^{a_2=b_1} a_{id} [bc]_{dj} = \sum_{d=1}^{a_2=b_1} a_{id} (\sum_{k=1}^{b_2=c_1} b_{dk} c_{kj}) = \sum_{1 \leq d \leq a_2, 1 \leq k \leq b_2} a_{id} b_{dk} c_{kj}$.
Hence the every corresponding entries of $(AB)C$ and $A(BC)$ are equal. Q.E.D.

3. Given a matrix P with non-zero determinant, and $\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{r_3}$ are 3×3 vectors such that $P = \begin{pmatrix} \boldsymbol{r_1}^T \\ \boldsymbol{r_2}^T \\ \boldsymbol{r_3}^T \end{pmatrix}$ (where T stands for transpose). Then $P^{-1} = \frac{1}{\det(P)} \begin{pmatrix} \boldsymbol{r_2} \times \boldsymbol{r_3} & \boldsymbol{r_3} \times \boldsymbol{r_1} & \boldsymbol{r_1} \times \boldsymbol{r_2} \end{pmatrix}$.

   Proof: Let $P^{-1} = \begin{pmatrix} \boldsymbol{s_1} \ \boldsymbol{s_2} \ \boldsymbol{s_3} \end{pmatrix}$. Observe that $\boldsymbol{r_i} \cdot \boldsymbol{s_j} = 0$ for $i \neq j$ and 1 otherwise. This means that $\boldsymbol{s_i}$ is a scalar multiple of $\boldsymbol{r_{i+1}} \times \boldsymbol{r_{i+2}}$ (indices taken modulo 3). (Why? $\boldsymbol{s_i}$ is perpendicular to both $\boldsymbol{r_{i+1}}$ and $\boldsymbol{r_{i+2}}$.) Now let constants k_i be that $\boldsymbol{s_i} = k_i(\boldsymbol{r_{i+1}} \times \boldsymbol{r_{i+2}})$, so $1 = k_i \cdot \boldsymbol{r_i} \cdot (\boldsymbol{r_{i+1}} \times \boldsymbol{r_{i+2}})$. However, it is well-known that $\boldsymbol{r_1} \cdot (\boldsymbol{r_2} \times \boldsymbol{r_3}) = \boldsymbol{r_2} \cdot (\boldsymbol{r_3} \times \boldsymbol{r_1}) = \boldsymbol{r_3} \cdot (\boldsymbol{r_1} \times \boldsymbol{r_2}) = \det(P)$ so $k_1 = k_2 = k_3 = \frac{1}{\det(P)}$ and we are done. To verify this for k_1, if $\boldsymbol{r_1} = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix}$, $\boldsymbol{r_2} = \begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \end{pmatrix}$, $\boldsymbol{r_3} = \begin{pmatrix} a_{31} \\ a_{32} \\ a_{33} \end{pmatrix}$ then $\boldsymbol{r_1} \cdot (\boldsymbol{r_2} \times \boldsymbol{r_3}) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{32}a_{21} = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \det(P)$.

4. If an $n \times n$ matrix Q has n linearly independent eigenvectors $\boldsymbol{r_1}, \boldsymbol{r_2}, \cdots \boldsymbol{r_n}$ and corresponding eigenvalues $a_1, a_2, \cdots, a_n$, then $Q = PDP^{-1}$, where $P = (\boldsymbol{r_1} \ \boldsymbol{r_2} \ \ldots \ \boldsymbol{r_n})$ and D is the diagonal matrix with a_i at i-th row and i-th column.

   Proof: Denote, again, T as the transpose of vector. If $P^{-1} = \begin{pmatrix} \boldsymbol{s_1}^T \\ \boldsymbol{s_2}^T \\ \vdots \\ \boldsymbol{s_n}^T \end{pmatrix}$ then $\boldsymbol{r_i} \cdot \boldsymbol{s_j} = 1$ if $i = j$ and 0 otherwise. Now let $n \times 1$ matrix x be $(\boldsymbol{r_i})$, then $P^{-1}x$ has 0 for all elements except the i-th entry, which is 1. This means $DP^{-1}x = D \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ has 0 for all elements except the i-th entry, which is a_i. Therefore $(\boldsymbol{r_1}\ \boldsymbol{r_2}\ \dots\ \boldsymbol{r_n})\begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_i(\boldsymbol{r_i}) = a_i x$. This means that PDP^{-1} has exactly the same eigenvectors (with corresponding eigenvalues) as Q, and we show that such matrix is unique.

To this end, let's show that $PDP^{-1}-Q \equiv 0$. Now we have $Q(r_i) = PDP^{-1}(r_i) = a_i(r_i)$, so $(PDP^{-1}-Q)(\boldsymbol{r_i}) = 0$ for $i = 1, 2, \cdots, n$. But since $\boldsymbol{r_1}, \boldsymbol{r_2}, \cdots, \boldsymbol{r_n}$ are linearly independent, we know that the dimension of null space of $PDP^{-1} - Q$ is n, or the rank of this matrix is 0. Hence $PDP^{-1} - Q$ is a zero matrix. Q.E.D.

5. Let A be an $m \times n$ matrix. If there exists a matrix B such that AB and BA are identity matrices, then $m = n$.

   Proof: If AB and BA are both defined, then the dimension of B is $n \times m$, and AB and BA have dimensions $m \times m$ and $n \times n$, respectively. If $m \neq n$, without the loss of generality we can assume that $m < n$. We prove that the matrix BA has rank of at most m, which forces its determinant to be zero (and therefore cannot be identity matrix).

   Now let $C = BA$, and denote its entry in vector manner $(\boldsymbol{v_1}\ \boldsymbol{v_2} \cdots \boldsymbol{v_n})$. Then $\boldsymbol{v_i} = \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix}$
   $= \begin{pmatrix} b_{11}a_{1i} + b_{12}a_{2i} + \cdots + b_{1m}a_{mi} \\ b_{21}a_{1i} + b_{22}a_{2i} + \cdots + b_{2m}a_{mi} \\ \vdots \\ b_{n1}a_{1i} + b_{n2}a_{2i} + \cdots + b_{nm}a_{mi} \end{pmatrix} = (a_{i1}\boldsymbol{w_1} + a_{i1}\boldsymbol{w_2} + \cdots + a_{im}\boldsymbol{w_m})$ where $\boldsymbol{w_j}$ represents the vector $\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}$, or the j-th column of matrix B. Notice that every vector $\boldsymbol{v_i}$ can be represented as $x_1\boldsymbol{w_1} + x_2\boldsymbol{w_2} + \cdots + x_m\boldsymbol{w_m}$, where $x_1, x_2, \cdots, x_n \in \mathbb{R}$. This means the set $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_n}\}$ can only span in at most m dimensions, i.e. the dimension of column space of this matrix BA is at most m.

6. $\det(A) \cdot \det(B) = \det(AB)$, where A, B are $n \times n$ matrices.
   Proof: Let $C = AB$, and denote a_{ij} as the entry of i-th row and j-th column of A. Define b_{ij} and c_{ij} similarly. The term $a_{pq}b_{rs}$ will appear as a term in C iff $q = r$, in which case it will be contained in the entry c_{ps}. Moreover, to be considered into the determinant of C, each term must be in the form of $\prod_{i=1}^{n} c_{i\sigma(i)} = \prod_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{j\sigma(i)}\right)$, where $\{\sigma(i) | 1 \leq i \leq n\}$ is the permutation of $\{1, 2, \cdots, n\}$. Individually speaking, the product of sums above comprises terms in the form $\prod_{i=1}^{n} a_{i\alpha(i)}b_{\alpha(i)\sigma(i)}$, where $\alpha(1), \alpha(2), \cdots \alpha(n)$ is a sequence of n numbers satisfying $1 \leq \alpha(j) \leq n$, $\forall j \in [1, n]$. Notice also that $\det(C) =$

$\sum \left(\prod_{i=1}^{n} (-1)^{N(\gamma)} c_{i\gamma(i)} \right)$, where he sum is taken across all permutations γ and $N(\gamma)$ is the number of inversions in γ. This means that the coefficient of $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ in $\det(C)$ is well-defined, i.e. the sums of coefficient of this term in all the terms $(-1)^{N(\gamma)} c_{i\gamma(i)}$. We show that if $\{\alpha(i) | 1 \leq i \leq n\}$ is not a permutation of $\{1, 2, \cdots, n\}$ (i.e. some of them are equal), then the product $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ has total coefficient zero in $\det(C)$. And if it is a permutation (meaning each term is different), then the coefficient is +1 if $\{\sigma(i) | 1 \leq i \leq n\}$ is an even permutation, and -1 otherwise.

Now for each particular permutation σ, the terms in the (multi-)set $\{a_{ij} b_{j\sigma(i)} | 1 \leq i, j \leq n\}$ is pairwise distinct, which means the term $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ has coefficient of exactly 1 in the expansion of $\prod_{i=1}^{n} c_{i\sigma(i)} = \prod_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{j\sigma(i)} \right)$. (I.e. every term in the form $)x = 1^n \prod a_{xy} b_{y\sigma(x)}$ ($\sigma=$ permutation of x) can only appear at most once in this expansion). Define the *occurences* of the term $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ as the total number of the permutation σ' such that this term is a term in the expansion of $\prod_{i=1}^{n} c_{i\sigma'(i)}$. Then for each k we name $f(k)$ as the number of $i \in [1, n]$ s.t. $\alpha(i) = k$. We show that the occurences of $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ is $\prod_{i=1}^{n} f(i)!$. Indeed, this is equivalent to saying that $\prod_{i=1}^{n} a_{i\alpha'(i)} b_{\alpha'(i)\sigma'(i)} = \prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$, or simply $\prod_{i=1}^{n} b_{\alpha(i)\sigma'(i)} = \prod_{i=1}^{n} b_{\alpha(i)\sigma(i)}$ as we can take $\alpha' \equiv \alpha$ for the purpose of establishing this proof of our claim. Now define S_k as $\{i | 1 \leq i \leq n, \alpha(i) = k\}$ (for $k = 1, 2, \cdots, n$) and what we need is $\{\sigma'(i) | i \in S(k)\} = \{\sigma(i) | i \in S(k)\}$, $\forall k \in [1, n]$. Now $\sigma' \equiv \sigma$ is definitely a solution, and for each such k we know that $\{\sigma(i) | i \in S(k)\}$ has $f(k)!$ solutions. For each k, there are $f(k)!$ possibilities to very σ, and multiplying this for all such k yields the answer $f(1)! f(2)! \cdots f(n)!$.

To finish this lemma, we need to prove that if $f(k) \geq 2$ for some k, then among all $\prod_{i=1}^{n} f(i)!$ permutations σ', exactly half of them is odd and half even. This means that there must be $\frac{\prod_{i=1}^{n} f(i)!}{2}$ occurences of $a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ that contibutes positively to the determinant of C (i.e. have coefficient +1 in this determinant calculation) and $\frac{\prod_{i=1}^{n} f(i)!}{2}$ that contributes negatively, which gives an overall coefficient of zero. We need the following well-known lemma (not hard to prove, though) :

If integer $x \geq 2$, swapping any two elements in a permutation of x distinct objects will change the permutation from odd to even, and vice versa.

This allows us to pair up all feasible permutations σ' into $\frac{\prod_{i=1}^{n} f(i)!}{2}$ pairs such that ech pair has an odd permutation and even permutation. Let j be an index in $[1, n]$ such that $f(j) \geq 2$, and choose j_1 and j_2 such that $\alpha(j_1) = \alpha(j_2) = j$. Then for each permutation σ', pair it up with another permutation σ_1' such that:

- $\sigma_1'(i_1) = \sigma'(i_2)$
- $\sigma_1'(i_2) = \sigma'(i_1)$
- $\sigma_1'(i) = \sigma'(i),\ \forall\ i \in [1,n]\backslash\{i_1, i_2\}$

Clearly, if σ' fulfills the claim that $\{\sigma'(i)|i \in S(k)\} = \{\sigma(i)|i \in S(k)\}$, $\forall k \in [1,n]$ then σ_1' also fulfills the claim that $\{\sigma_1'(i)|i \in S(k)\} = \{\sigma(i)|i \in S(k)\}$, $\forall k \in [1,n]$. Moreover, each pair is disjoint (meaning no element can exist in two groups) since the other element (i.e. permutation) in the same pair with a permutation is defined uniquely before. This completes our bijection (and hence our claim).

If $f(k) = 1$ for all k, then there is only one such occurence of $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$, or $\sigma' \equiv \sigma$ is the only possible permutation. The two previous terms have coefficient 1 in $\det(C)$ iff $\prod_{i=1}^{n} c_{i\sigma(i)}$ has a coefficient 1 in $\det(C)$, i.e. σ is an even permutation or $N(\sigma)$ is even.

Finally, let's prove that for if α is a permutation, then the coefficient of $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ in $\det(C)$ is the product of the coefficient of $\prod_{i=1}^{n} a_{i\alpha(i)}$ in $\det(A)$ and the coefficient of $\prod_{i=1}^{n} b_{\alpha(i)\sigma(i)}$ n $\det(B)$. Let I be the identity permutation, i.e. $I(x) = x$, $\forall x \in [1,n]$. Denote also the permutation that brings α to σ as ω (means if $\alpha(i) = j$ and $\sigma(i) = k$, then $\omega(j) = k$. Try to think ω as $\sigma\alpha^{-1}$.) Now we know that if a permutation σ is even, then I can be mapped to σ by swapping two neighbouring elements only by an even number of times (and vice versa). In other words, if N denotes the minimum number of swappings needed, then we can map I to σ by $N(\alpha) + N(\omega)$ of neighbouring swappings (first $N(\alpha)$ swapping from I to α, then $N(\omega)$ swappings from α to σ). This entails $N(\sigma) \equiv N(\alpha) + N(\omega) \pmod 2$, or $(-1)^{N(\sigma)} = (-1)^{N(\alpha)} \cdot (-1)^{N(\omega)}$. Since $\prod_{i=1}^{n} b_{\alpha(i)\sigma(i)} = \prod_{i=1}^{n} b_{i\omega(i)}$, we have: coefficient of $\prod_{i=1}^{n} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ is $(-1)^{N(\sigma)}$, coefficient of $\prod_{i=1}^{n} a_{i\alpha(i)}$ is $(-1)^{N(\alpha)}$, and $\prod_{i=1}^{n} b_{\alpha(i)\sigma(i)}$ is $(-1)^{N(\omega)}$. This proves our claim, and $\det(C) = \prod_{i=1}^{n} (-1)^{N(\sigma)} c_{i\sigma(i)} = \sum \prod_{i=1}^{n} (-1)^{N(\sigma)} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ (σ any permutation, α any n-tuple) $= \sum \prod_{i=1}^{n} (-1)^{N(\sigma)} a_{i\alpha(i)} b_{\alpha(i)\sigma(i)}$ (same definition for σ but restrict α to permutations, since the other case has coefficient zero and can be disregarded)$= \sum (-1)^{N(\alpha)} a_{i\alpha(i)} \cdot \sum (-1)^{N(\omega)} b_{i\omega(i)} = \sum (-1)^{N(\alpha)} a_{i\alpha(i)} \cdot \sum (-1)^{N(\omega)} b_{\alpha(i)\sigma(i)} = \det(A)\det(B)$, as desired.

4 Centroid of bodies.

1. Solid hemisphere with radius r.
   Answer: $\frac{3}{8}r$ from the centre, and lying on the line through centre and perpendicular to the plane of the great circle.
   Proof: Let the radius of the hemisphere be r, then the equation of this hemisphere is just $y = \sqrt{r^2 - x^2}$ from 0 to r having rotated 360°. We know that $\bar{y} = 0$. To find $\bar{x}$ simply

take the fraction $\dfrac{\pi\int_0^r xy^2\,dx}{\pi\int_0^r y^2\,dx}=\dfrac{\int_0^r x(r^2-x^2)\,dx}{\int_0^r r^2-x^2\,dx}=\dfrac{[\frac{x^2r^2}{2}-\frac{x^4}{4}]_0^r}{[xr^2-\frac{x^3}{3}]_0^r}=\dfrac{\frac{r^4}{4}}{\frac{2r^3}{3}}=\frac{3}{8}r.$

2. Hemispherical shell with radius r.
Answer: $\frac{1}{2}r$ from the centre, and lying on the line through centre and perpendicular to the plane of the great circle.
Proof: same equation as above for x and y, but now we need $\bar{x}=\dfrac{2\pi\int_0^r xy\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}\,dx}{2\pi\int_0^r y\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}\,dx}$.
Observe that $\dfrac{d}{dx}((r^2-x^2)^{0.5})=\dfrac{-x}{(r^2-x^2)^{0.5}}$ so $y\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}=(r^2-x^2)^{0.5}\cdot\sqrt{r^2(r^2-x^2)^{-1}}=r$. The original integral then becomes $\dfrac{\int_0^r xr\,dx}{\int_0^r r\,dx}=\dfrac{[r(\frac{x^2}{2})]_0^r}{[r(x)]_0^r}=\dfrac{\frac{r^3}{2}}{r^2}=\dfrac{1}{2}r.$

3. Circular sector with radius r and subtended angle 2α.
Answer: $\dfrac{2r\sin\alpha}{3\alpha}$ from the centre of the circle (i.e. sector) and lying on the angle bisector of the sector.
Proof: Same equation as above, for x from $r\cos 2\alpha$ to r, and $y=\tan 2\alpha$ for x from 0 to $r\cos 2\alpha$. Now $\bar{x}=\dfrac{\int_0^r xy\,dx}{\int_0^r y\,dx}=\dfrac{\int_0^{r\cos 2\alpha} x^2\tan 2\alpha\,dx+\int_{r\cos 2\alpha}^r x\sqrt{r^2-x^2}\,dx}{\int_0^{r\cos 2\alpha} x\tan 2\alpha\,dx+\int_{r\cos 2\alpha}^r \sqrt{r^2-x^2}\,dx}$. Now to deal with the second integral in numerator we need the substitution $x=r\cos\theta$, and we have $dx=-r\sin\theta d\theta$. $\int_{r\cos 2\alpha}^r x\sqrt{r^2-x^2}\,dx=\int_{2\alpha}^0 -r\cos\theta\sqrt{r^2-r^2\cos^2\theta}r\sin\theta d\theta=\int_0^{2\alpha} r^3\cos\theta\sin^2\theta d\theta=\left[r^3\dfrac{\sin^3\theta}{3}\right]_0^{2\alpha}=\dfrac{r^3\sin^3 2\alpha}{3}$. (We know from the area of sector that the denominator is $r^2\alpha$.) $\int_0^{r\cos 2\alpha} x^2\tan 2\alpha\,dx=\left[\dfrac{x^3\tan 2\alpha}{3}\right]_0^{r\cos 2\alpha}=\dfrac{(r\cos 2\alpha)^3\sin 2\alpha}{3\cos 2\alpha}=\dfrac{r^3\sin 2\alpha\cos^2 2\alpha}{3}$. Adding the two terms up yield $\dfrac{r^3\sin 2\alpha}{3}(\cos^2 2\alpha+\sin^2 2\alpha)=\dfrac{r^3\sin 2\alpha}{3}$.
This yields $\bar{x}=\dfrac{r^3\sin 2\alpha}{3r^2\alpha}=\dfrac{r\sin 2\alpha}{3\alpha}$.
It is obvious that the centre of mass of this sector lies on the angle bisector of the sector due to symmetry, hence have gradient $\cos\alpha$ rom centre O (the origin). The length from the origin to the centre of mass has therefore length $\dfrac{r\sin 2\alpha}{3\alpha\cos\alpha}=\dfrac{2r\sin\alpha}{3\alpha}$.

4. Circular arc with radius r and subtended angle 2α.
Answer: $\dfrac{r\sin\alpha}{\alpha}$ from the centre of the circle (i.e. sector) and lying on the angle bisector of the sector.
Proof: The equation is given by $y=\sqrt{r^2-x^2}$, x from rcos$\,2\alpha$ to r. As shown above, $\sqrt{1+\left(\frac{dy}{dx}\right)^2}=\frac{r}{\sqrt{r^2-x^2}}$. Here, $\bar{x}=\dfrac{\int_{r\cos 2\alpha}^r \frac{xr}{\sqrt{r^2-x^2}}\,dx}{\int_{r\cos 2\alpha}^r \frac{r}{\sqrt{r^2-x^2}}\,dx}$. The denominator is simply the arc length $2r\alpha$, while the numerator is equal to $[-r\sqrt{r^2-x^2}]_{r\cos 2\alpha}^r=r\sqrt{r^2-r^2\cos^2 2\alpha}=r^2\sin 2\alpha$. Therefore $\bar{x}=\dfrac{r^2\sin 2\alpha}{2r\alpha}$. Again, divide it by $\cos\alpha$ to find he distance from the centre and we have $\dfrac{r^2\sin 2\alpha}{2r\alpha\cos\alpha}=\dfrac{r\sin\alpha}{\alpha}$.

5. Triangular lamina ABC.
Answer: Let the midpoint of BC to be M, then the centre of mass is the point G on segment AM, satisfying $AG : GM = 2 : 1$.
Proof: We prove that this centroid must lie on all medians of a triangle by establishing that the magnitude of moment with respect to AM at both sides of AM must be the same. To do this, we cut the triangle into very tiny strips, each one parallel to AM, and prove that for each strip, we can find another strip on the other side of the median that has the same length (or same mass) and with the same distance from the line AM.

We need to prove first that the number of strips on both sides must the same, given the each has width d (infinisimal). Now, the product of d and the total number of strips on side BM is the perpendicular distance from B to AM, which is $BM \sin\angle AMB$; the product of d and the total number of strips on side CM is the perpendicular distance from C to AM, which is $CM \sin\angle AMC$. However, $BM = CM$ (M is the midpoint of BC) and $\sin\angle AMB = \sin\angle AMC$ (these angles are supplementary) so this claim is proven.

Next, consider a point P on side BM, and Q a point on side AM such that PQ is parallel to AM. Next, reflect P across M to get P', and let Q' be a point on AC such that $P'Q' \parallel AM$. Now $\dfrac{PQ}{AM} = \dfrac{BP}{BM} = \dfrac{CP'}{CM} = \dfrac{P'Q'}{AM}$ so $PQ = P'Q'$. (The first equivalence is because triangles BQP and BAM are similar; the second equivalence follows when $BP = CP'$ by the definition of reflection and $BM = CM$; the third equivalence is the consequence of the fact that triangles CAM and $CQ'P'$ are similar). Finally, the distance between strip PQ and line AM is $PM \sin\angle BAM$ and distance between strip $P'Q'$ and line AM is $PM \sin\angle CAM$. It is easy to verify that the are the same.

The similar proof above can be used to verify that the other medians contain the centre of mass of triangle ABC as well.

5 Moment of inertia.

1. Thin rod of mass m length $2r$ about the perpendicular axis through the midpont.
Answer: $\dfrac{mr^2}{3}$.
Proof: Moment of inertia$=m \cdot \dfrac{\int_{-r}^{r} x^2\,dx}{\int_{-r}^{r} 1\,dx} = m \cdot \dfrac{[\frac{x^3}{3}]_{-r}^{r}}{[x]_{-r}^{r}} = m \cdot \dfrac{\frac{2r^3}{3}}{2r} = \dfrac{mr^2}{3}$.

2. Rectangular lamina of mass m and dimensions $2a \times 2b$ about the perpendicular axis through the centre of mass.
Answer: $\dfrac{1}{3}m(a^2 + b^2)$.
Proof: Consider al the particles on the rectangular lamina (with mass dm), so the mass of the particles is actually $\sum\limits_{-a\le x\le a,-b\le y\le b} dm$ and the moment of inertia is $\sum\limits_{-a\le x\le a,-b\le y\le b} (x^2 + y^2)dm$. Turning infinitely many particles into double integration (because there are two dimensions!) we have $\int_{-a}^{a}(\int_{-b}^{b} 1\,dy)\,dx$ for mass, and $\int_{-a}^{a}(\int_{-b}^{b} x^2 + y^2\,dy)\,dx$ for moment of inertia. Now $(\int_{-b}^{b} 1\,dy)=[y]_{-b}^{b} = 2b$ so $\int_{-a}^{a}(\int_{-b}^{b} 1\,dy)\,dx = \int_{-a}^{a} 2b\,dx = [2bx]_{-a}^{a} = 4ab$. $(\int_{-b}^{b} x^2 + y^2\,dy)=[x^2y + \frac{y^3}{3}]_{-b}^{b} = 2x^2b + \frac{2b^3}{3}$ so $\int_{-a}^{a} 2x^2b + \frac{2b^3}{3}\,dx = [\dfrac{2x^3b + 2xb^3}{3}]_{-a}^{a} = \dfrac{4a^3b + 4ab^3}{3} = 4ab\left(\dfrac{a^2 + b^2}{3}\right)$. Dividing it by $4ab$ and multiplying by m yields the moment

of inertia as $m\left(\dfrac{a^2+b^2}{3}\right)$.

3. Disc/solid cylinder with mass m and radius r about the axis pependicular to the circular plane and pasing through its centre.

   Answer: $\dfrac{mr^2}{2}$.

   Proof: Let the height of solid cylinder be h (in the event of a circular lamina we can treat this h as infinisimal (which doesn't affect our answer anyway). Now consider the sub-ring (or sub-cylinderic shell) with centre the axis and radius x, and the surface area is definitely $2\pi xh$. The desired ratio now becomes $\dfrac{\int_0^r 2\pi x^3 h\,dx}{\int_0^r 2\pi xh\,dx} = \dfrac{\int_0^r x^3\,dx}{\int_0^r x\,dx} = \dfrac{[x^4/4]_0^r}{[x^2/2]_0^r} = \dfrac{r^2}{2}$, which leads to our answer.

4. Solid sphere of mass m and radius r about an axis passing through its centre.

   Answer: $\dfrac{2mr^2}{5}$.

   Proof: Treat this sphere as $x^2+y^2+z^2 \le r^2$ about the z-axis, and consider the locus of points of distance d from this z-axis. Now $d = x^2+y^2$, and $d+z^2 \le r^2$. Therefore $|z| \le \sqrt{r^2-d^2}$ or $-\sqrt{r^2-d^2} \le z \le \sqrt{r^2-d^2}$. The locus is therefore the curved face of cylindrical shell with radius d and height $2\sqrt{r^2-d^2}$, and the surace area is $2\pi d(2\sqrt{r^2-d^2}) = 4\pi d\sqrt{r^2-d^2}$.

   The desired fraction is now $\dfrac{\int_0^r (x^2)4\pi x\sqrt{r^2-x^2}\,dx}{\int_0^r 4\pi x\sqrt{r^2-x^2}\,dx} = \dfrac{\int_0^r x^3\sqrt{r^2-x^2}\,dx}{\int_0^r x\sqrt{r^2-x^2}\,dx}$. Now for both denominator and numerator, we use the substitution $x = r\sin\theta$ to get $dx = r\cos d\theta$, so that $\int_0^r x\sqrt{r^2-x^2}\,dx = \int_0^{\frac{\pi}{2}}(r\sin\theta)(r\cos\theta)(r\cos\theta\,d\theta) = \int_0^{\frac{\pi}{2}} r^3\sin\theta\cos^2\theta\,d\theta = r^3\left[\dfrac{-\cos^3\theta}{3}\right]_0^{\frac{\pi}{2}} = \dfrac{r^3}{3}$.

   Similarly, $\int_0^r x^3\sqrt{r^2-x^2}\,dx = \int_0^{\frac{\pi}{2}}(r\sin\theta)^2(r\sin\theta)(r\cos\theta)(r\cos\theta\,d\theta) = \int_0^{\frac{\pi}{2}} r^5\sin^3\theta\cos^2\theta\,d\theta$ $= \int_0^{\frac{\pi}{2}} r^5\sin\theta(1-\cos^2\theta)\cos^2\theta\,d\theta = \int_0^{\frac{\pi}{2}} r^5\sin\theta\cos^2\theta\,d\theta$ - $\int_0^{\frac{\pi}{2}} r^5\sin\theta\cos^4\theta\,d\theta = r^5\left[\dfrac{-\cos^3\theta}{3}\right]_0^{\frac{\pi}{2}}$ - $r^5\left[\dfrac{-\cos^5\theta}{5}\right]_0^{\frac{\pi}{2}} = \dfrac{r^5}{5} - \dfrac{r^5}{3} = \dfrac{2r^5}{15}$. Summarizing above yields the desired fraction as $\dfrac{2r^5}{15} \div \dfrac{r^3}{3} = \dfrac{2r^2}{5}$, whch yields the answer.

5. Spherical shell of mass m and radius r about an axis passing through its centre.

   Answer: $\dfrac{2mr^2}{3}$.

   Proof: Treat this shell as the equation $y = \sqrt{r^2-x^2}$, rotated through 360° the x-axis. (x rnging from $-r$ to r). Assume, too, that the x-axis is our axis of reference. Now the distance of each point from this axis is its y- coordinate, so for each y, $x = \pm\sqrt{r^2-x^2}$.

   For each of these two x, a circle with circumference $2\pi y\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}$ is generated upon reolution (recall how we found surface area of revolution of a graph!) As always, $1+\left(\dfrac{dy}{dx}\right)^2 = \dfrac{r^2}{r^2-x^2}$. This means $2\pi y\sqrt{1+\left(\dfrac{dy}{dx}\right)^2} = 2\pi\sqrt{r^2-x^2}\dfrac{r}{\sqrt{r^2-x^2}} = 2\pi r$. This means

that our desired fraction is now $\dfrac{\int_{-r}^{r}(y^2)2\pi y\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}\,dx}{\int_{-r}^{r}2\pi y\sqrt{1+\left(\dfrac{dy}{dx}\right)^2}\,dx} = \dfrac{\int_{-r}^{r}2\pi r(r^2-x^2)\,dx}{\int_{-r}^{r}2\pi r\,dx} =$
$\dfrac{[2\pi r^3 x - \frac{2rx^3}{3}]_{-r}^{r}}{[2\pi rx]_{-r}^{r}} = \dfrac{4\pi r^4 - \frac{4}{3}r^4}{4\pi r^2} = \dfrac{2r^2}{3}$ so the answer is $\dfrac{2mr^2}{3}$.

6. **Perpendicular axis theorem.** (applies only to a lamina)
Proof: We assume that the lamina is on the $x-y$ plane, and Let $f(x,y)$ be the density at point x, y. Now the mass of this lamina is (ignoring limits) $\int(\int f(x,y)\,dx)\,dy$. For each particle at point (x,y), the distance from x-axis is y and the distance from y-axis is x. Its distance from z-axis is $\sqrt{x^2+y^2}$ (i.e. distance from the origin). Therefore $I_z = \int(\int(x^2+y^2)f(x,y)\,dx)\,dy = \int(\int x^2 f(x,y)\,dx + \int y^2 f(x,y)\,dx)\,dy = \int(\int x^2 f(x,y)\,dx)dy + \int(\int y^2 f(x,y)\,dx)\,dy = I_y + I_x$, as desired.

7. **Parallel axis theorem.** (applies to a solid and a lamina)
Proof: Let the density at point (x,y,z) be $f(x,y,z)$, and its mass is $m = \int(\int(\int f(x,y,z)\,dx)\,dy)\,dz$ (again, dropping limits). Assume, W.L.O.G., that the centre of mass lies at point (0,0,0), meaning that $\int(\int(\int xf(x,y,z)\,dx)\,dy)\,dz = \int(\int(\int yf(x,y,z)\,dx)\,dy)\,dz = \int(\int(\int zf(x,y,z)\,dx)\,dy)\,dz = 0$. Now let the axis through the centre of mass to be the z-axis, and the other axis to be $x = a, y = b$. The distance of solid from the first axis is $\sqrt{x^2+y^2}$ and distance from the second axis is $\sqrt{(x-a)^2+(y-b)^2}$. Now the moment of inertia about the second axis is $\int(\int(\int((x-a)^2+(y-b)^2)f(x,y,z)\,dx)\,dy)\,dz = \int(\int(\int(x^2+y^2+a^2+b^2-2ax-2by)f(x,y,z)\,dx)\,dy)\,dz = \int(\int(\int(x^2+y^2)f(x,y,z)\,dx)\,dy)\,dz + \int(\int(\int(a^2+b^2)f(x,y,z)\,dx)\,dy)\,dz$ - $2\int(\int(\int axf(x,y,z)\,dx)\,dy)\,dz$ - $2\int(\int(\int byf(x,y,z)\,dx)\,dy)\,dz = I_{z-\text{axis}} + (a^2+b^2)\int(\int(\int f(x,y,z)\,dx)\,dy)\,dz$ - $2a\int(\int(\int xf(x,y,z)\,dx)\,dy)\,dz$ - $2b\int(\int(\int yf(x,y,z)\,dx)\,dy)\,dz = I_{z-\text{axis}} + (a^2+b^2)m$ - 0, i.e. the moment of inertia about the first axis + mr^2 where r is the distance of the second axis from the origin.

6 Regression lines

1. Given n points (x_i, y_i), $i = 1, 2, \cdots, n$. Then the equation of line of regression of y on x is given by $y - \bar{y} = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}(x-\bar{x})$.
Proof: First, we prove that for a set of lines with common gradient, the sum of square of vertical deviation of each point from the line is minimum when the line passes through the centre of mass of the points.
Suppose that the gradient is k, and let $\theta = \tan^{-1}k$ be the angle (anticlockwise) of the line with the x-axis. Now perform a rotation of everything (the n points and the line) of angle θ clockwise about any point, say, the origin. Then the line now becomes horizontal, the initial vertical deviation of each point to the line becomes the "θ -degree to the vertical" deviation to the line (which is, $|\frac{1}{\cos\theta}|$ times the new vertical deviation (or perpendicular distance) to the line: multipliation by a constant across all points). Moreover, the centre of mass of the points also follows the same rotation. (It makes prefect sense intuitively, but to rigorize the argument, notice that $\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$ rotates the point (x,y) an-

gle α anticlockwise. If $\alpha = -\theta$ then the rotation of points (x_i, y_i), $i = 1, 2, \cdots, n$ have centre of mass $\frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \frac{1}{n}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\sum_{i=1}^{n}\begin{pmatrix} x_i \\ y_i \end{pmatrix}$, which is the original centre of mass having rotated angle α).

Now let (x_i', y_i') be our new points, and since the line is now horizontal we can assume that its equation is $y = w$, w constant. Now the vertical distance of each point to the line is $|y_i' - w|$ so the sum of square of distance is $\sum_{i=1}^{n}(y_i' - w)^2 = nw^2 - 2w\sum_{i=1}^{n} y_i' + \sum_{i}^{n}(y_i')^2 = n\left(w - \frac{\sum_{i=1}^{n} y_i'}{n}\right)^2 - \frac{\left(\sum_{i=1}^{n} y_i'\right)^2}{n} + \sum_{i}^{n}(y_i')^2$. Now the "$\theta$ degree to the normal" distance can be obtained by dividing each term by $|\cos\theta|$, which is a costant across each term! Therefore, the minimum distance can be achieved when $w = \frac{\sum_{i=1}^{n} y_i'}{n}$, i.e. passing through the centre of mass.

To find the gradient of this line, notice that the equation of the line must be in the form $y - \bar{y} = b(x - \bar{x})$. The vertical deviation from each point to the line is thus $y_i - b(x_i - \bar{x}) - \bar{y}$. Therefore the sum of square of vertical deviation is $\sum_{i=1}^{n}[(y_i - \bar{y}) - b(x_i - \bar{x})]^2 = \sum_{i=1}^{n} b^2(x_i - \bar{x})^2 - 2b\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) + \sum_{i=1}^{n}(y_i - \bar{y})^2$. Again, by completing the square we know that the least square sum is achieved when $b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$. Notice, also, that the numerator is equivalent to $\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}$ and that the denominator is equivalent to $\sum_{i=1}^{n}(x_i^2) - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$.