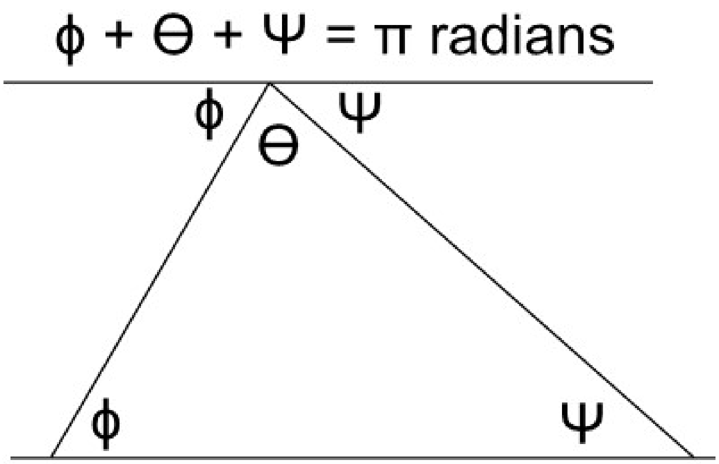
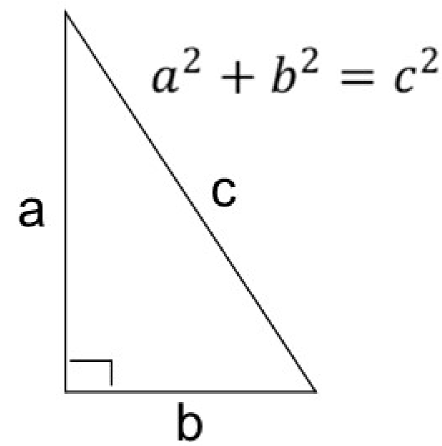
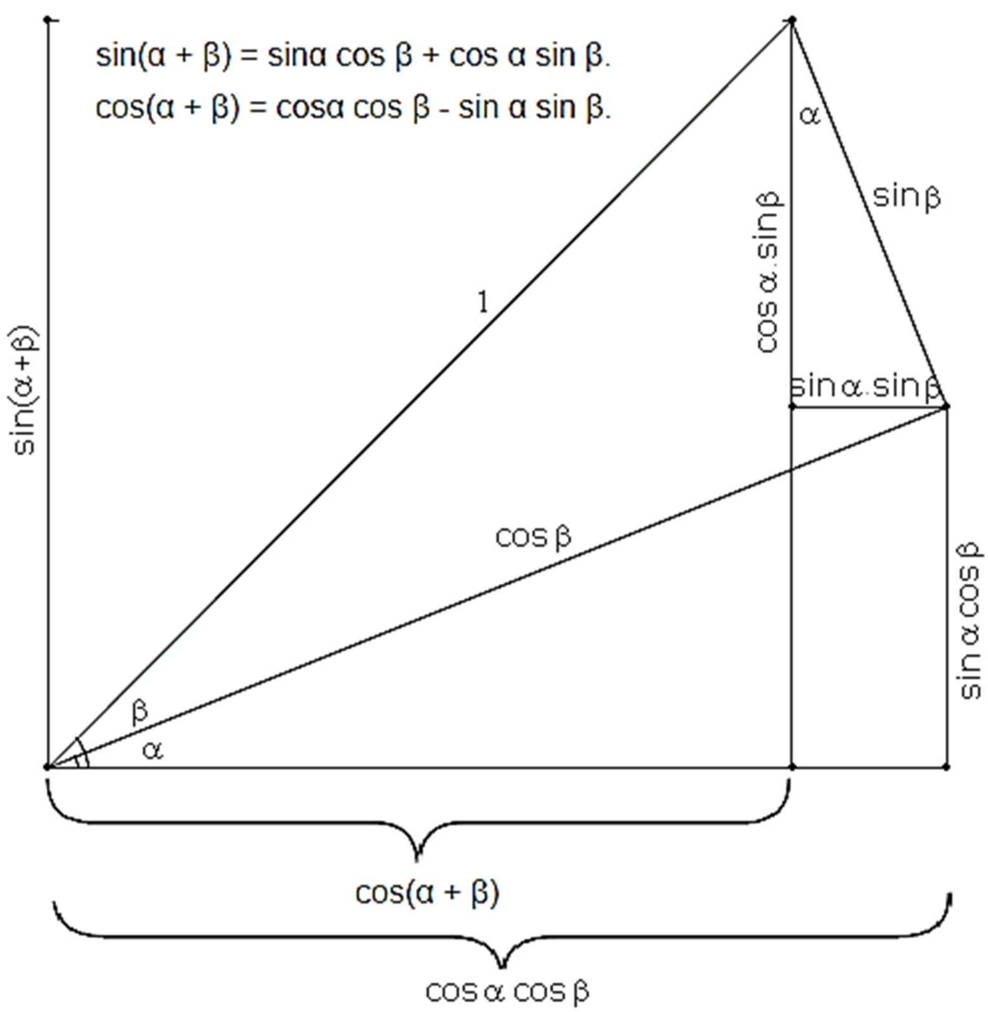
Derivation of the QM ≥ AM ≥ GM ≥ HM relationship from Axioms:

AXIOM 1 & AXIOM 2: Parallel Line Postulate of Euclid; the sum of the three internal angles of a triangle is equal to π radians (180 degrees). This is an alternative statement of Euclid's parallel line postulate. Note that the Pythagorean Theorem extends to an arbitrary number of dimensions with many proofs available.

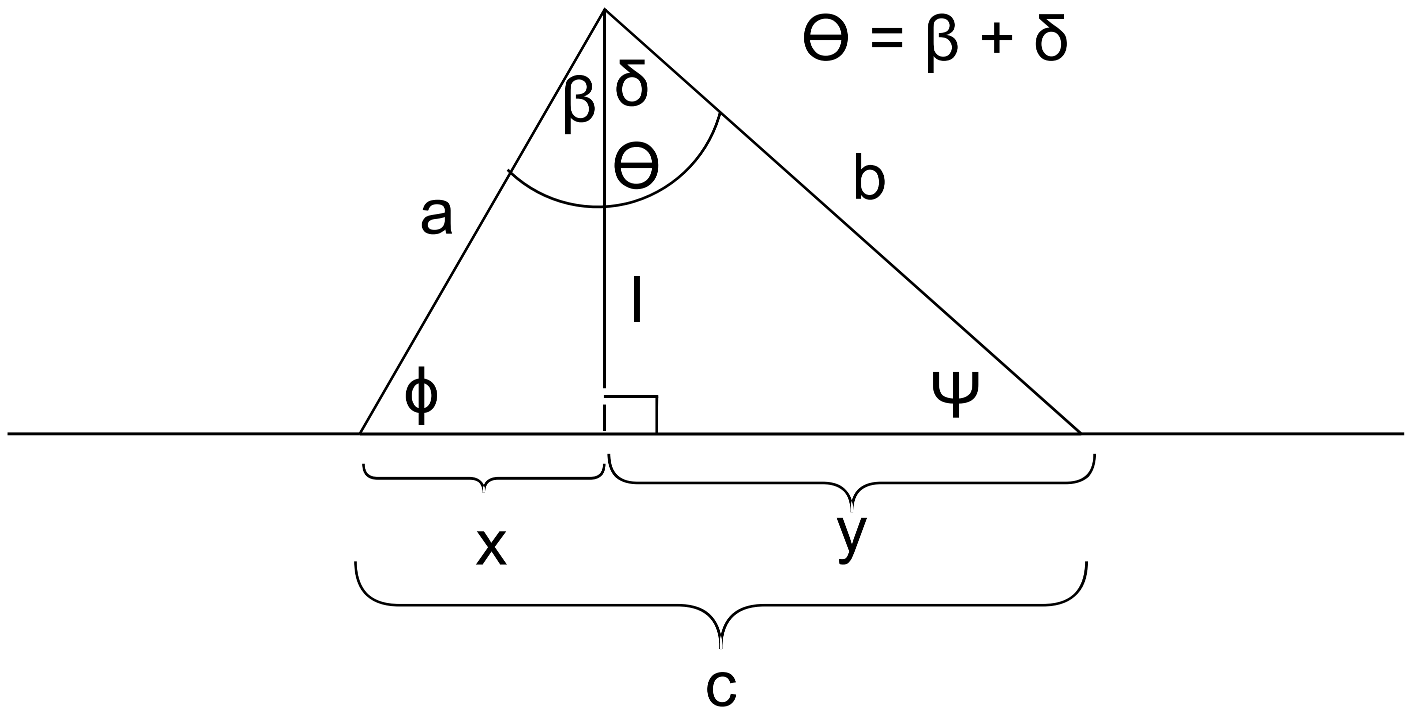




Derivation of cosine of sum of angles by AXIOM 1 (sum internal ∠'s of Δ = 180 degrees):



Derivation of the Law of Cosines from the Pythagorean Theorem:



$$x^2 + l^2 = a^2 \qquad y^2 + l^2 = b^2$$

$$a^2 + b^2 = 2l^2 + x^2 + y^2$$

$$c^2 = (x + y)^2 = x^2 + y^2 + 2xy$$

$$c^2 - 2xy = x^2 + y^2$$

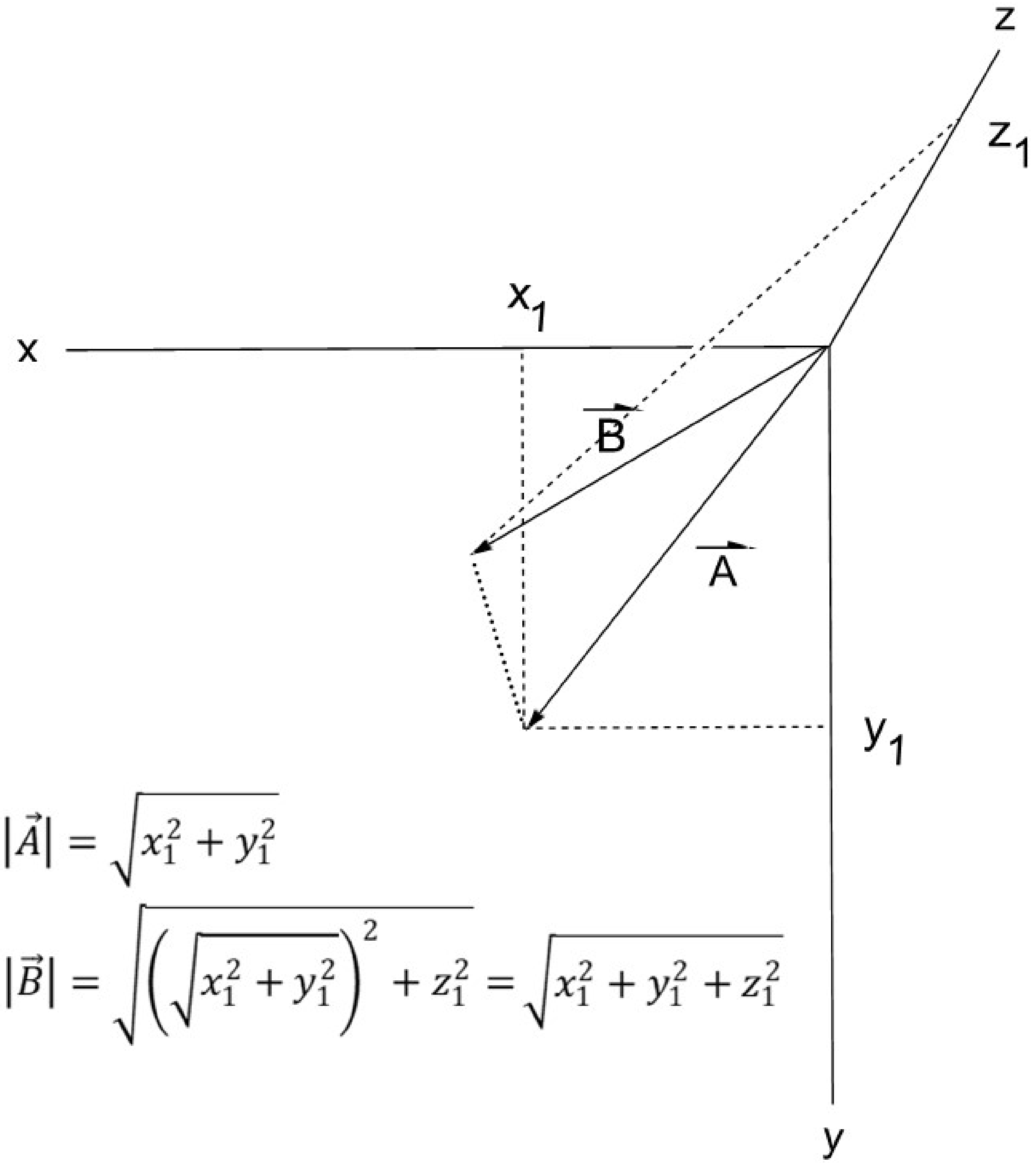
$$a^2 + b^2 = c^2 + 2l^2 - 2xy$$

$$l = acos(\beta) = bcos(\delta) \qquad x = \ asin(\beta) = bsin(\delta)$$

$$a^2 + b^2 = c^2 + 2abcos(\beta)\cos{(\delta)} - 2absin(\beta)\sin{(\delta)}$$

$$a^2 + b^2 = c^2 + 2abcos(\beta + \delta) = c^2 + 2abcos(\theta)$$

Derivation of the length of an “n” dimensional vector:



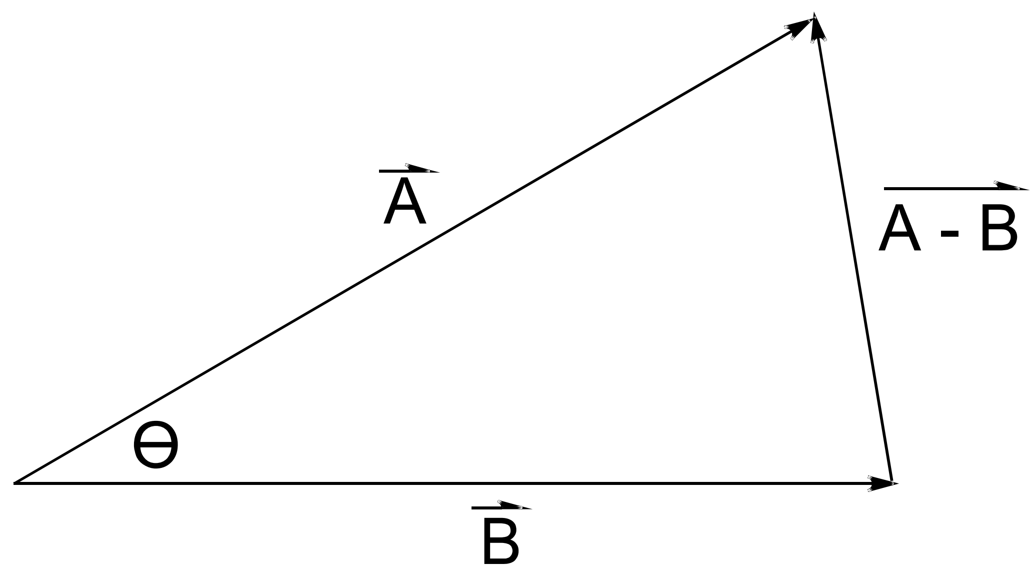
$$|\vec{A}| = \sqrt{x_1^2 + y_1^2}$$

$$|\vec{B}| = \sqrt{\left(\sqrt{x_1^2 + y_1^2}\right)^2 + z_1^2} = \sqrt{x_1^2 + y_1^2 + z_1^2}$$

The length of a vector in 2D (*e.g.*, vector A) is given by the Pythagorean Theorem. If one extends a 2D vector into 3D space (*e.g.*, vector B), one gets the new length by another application of the Pythagorean Theorem. Now imagine the 3D vector moved into a 2D space obtained by rotating the coordinate system and again extending this formerly 3D vector into a new dimension “d” (effectively a 4D vector). The new length is obtained by another application of the Pythagorean Theorem, and the process is continued on to “n” dimensions.

$$|\vec{V}| = \sqrt{x_1^2 + y_1^2 + z_1^2 + d_1^2 + \cdots + n_1^2}$$

Derivation of the Cauchy-Schwarz Inequality via Pythagorean Theorem & The Law of Cosines:



Consider n-dimensional vectors:

$$\vec{A} = x_a + y_a + z_a + \cdots + n_a$$

$$\vec{B} = x_b + y_b + z_b + \cdots + n_b$$

$$\overrightarrow{A - B} = (x_a - x_b) + (y_a - y_b) + (z_a - z_b) + \cdots + (n_a - n_b)$$

$$|\bar{A}| = \sqrt{x_a^2 + y_a^2 + z_a^2 + \cdots + n_a^2}$$

$$|\bar{B}| = \sqrt{x_b^2 + y_b^2 + z_b^2 + \cdots + n_b^2}$$

$$|\overline{A - B}| = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2 + \cdots + (n_a - n_b)^2}$$

Apply The Law of Cosines:

$$|\bar{A}|^2 + |\bar{B}|^2 = |\overline{A - B}|^2 + 2|\bar{A}||\bar{B}|\cos(\theta)$$

$$2x_a x_b + 2y_a y_b + 2z_a z_b + \cdots + 2n_a n_b = 2|\bar{A}||\bar{B}|\cos(\theta)$$

$$\vec{A} \cdot \vec{B} = x_a x_b + y_a y_b + z_a z_b + \cdots + n_a n_b = |\bar{A}||\bar{B}|\cos(\theta)$$

$$|\bar{A}||\bar{B}| \geq \vec{A} \cdot \vec{B} = x_a x_b + y_a y_b + z_a z_b + \cdots + n_a n_b$$

Proof that Quadratic Mean ≥ Arithmetic Mean via The Cauchy-Schwarz Inequality:

$$data = \{x_1, x_2, x_3, \dots, x_n\} = \text{"n"}\ dimensional\ vector$$

$$Quadtric\ Mean = QM = \frac{\sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}}{n}$$

$$Arithmetic\ Mean = AM = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

Consider the dot product of an “n” dimensional vector of all ones with our data vector:

$$\{x_1, x_2, x_3, \dots, x_n\} \cdot \{1, 1, 1, \dots, 1\} = x_1 + x_2 + x_3 + \cdots + x_n$$

$$|\{1, 1, 1, \dots, 1\}| = \sqrt{n}$$

$$|\{x_1, x_2, x_3, \dots, x_n\}| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}$$

$$\sqrt{n}\sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2} \geq x_1 + x_2 + x_3 + \cdots + x_n$$

$$n(x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2) \geq (x_1 + x_2 + x_3 + \cdots + x_n)^2$$

$$\frac{n(x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2)}{n^2} \geq \frac{(x_1 + x_2 + x_3 + \cdots + x_n)^2}{n^2}$$

$$\sqrt{\frac{(x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2)}{n}} \geq \sqrt{\frac{(x_1 + x_2 + x_3 + \cdots + x_n)^2}{n^2}}$$

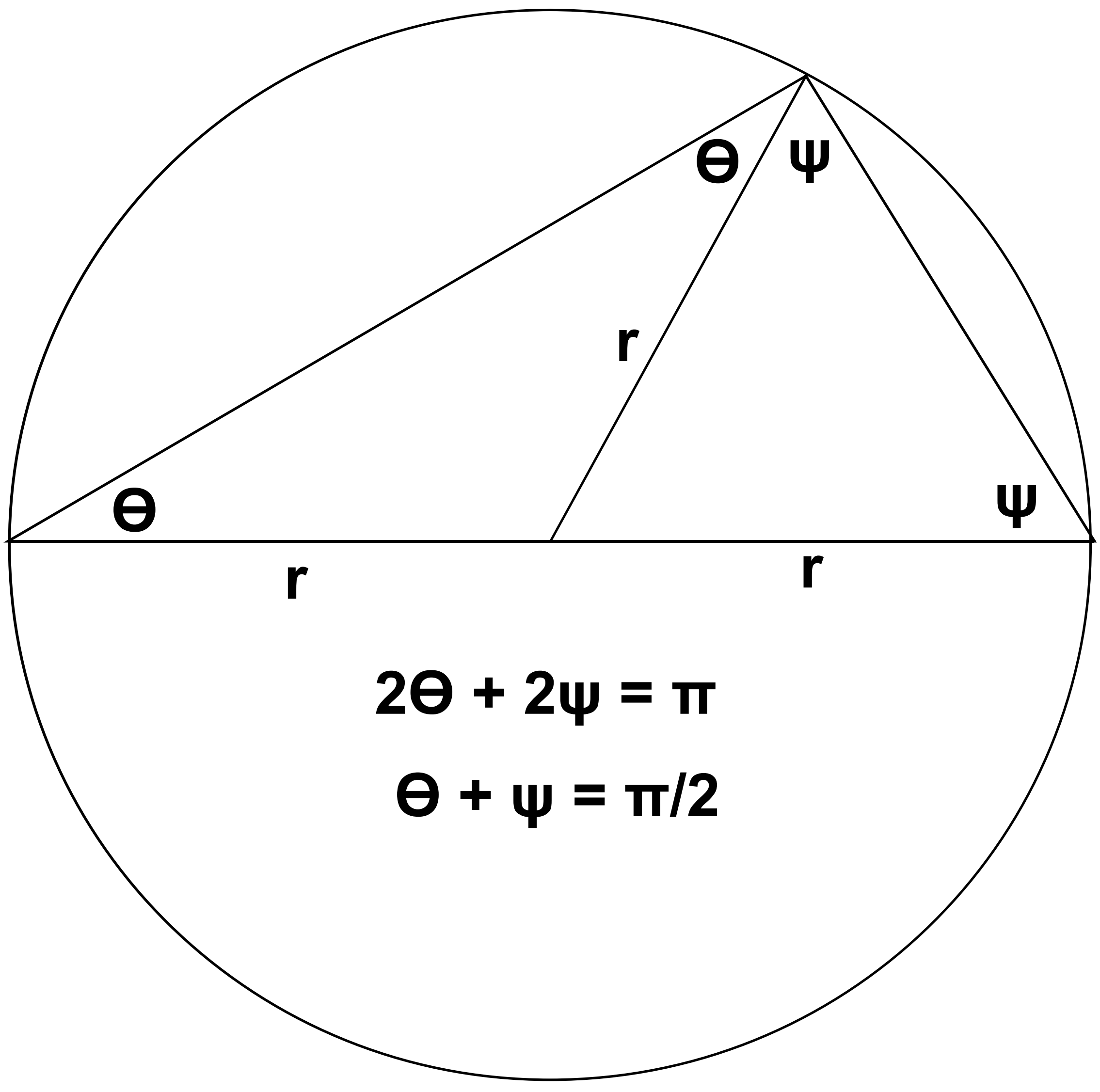
$$\sqrt{\frac{(x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2)}{n}} \geq \frac{(x_1 + x_2 + x_3 + \cdots + x_n)}{n}$$

$$QM = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \geq \frac{\sum_{i=1}^{n} x_i}{n} = AM$$

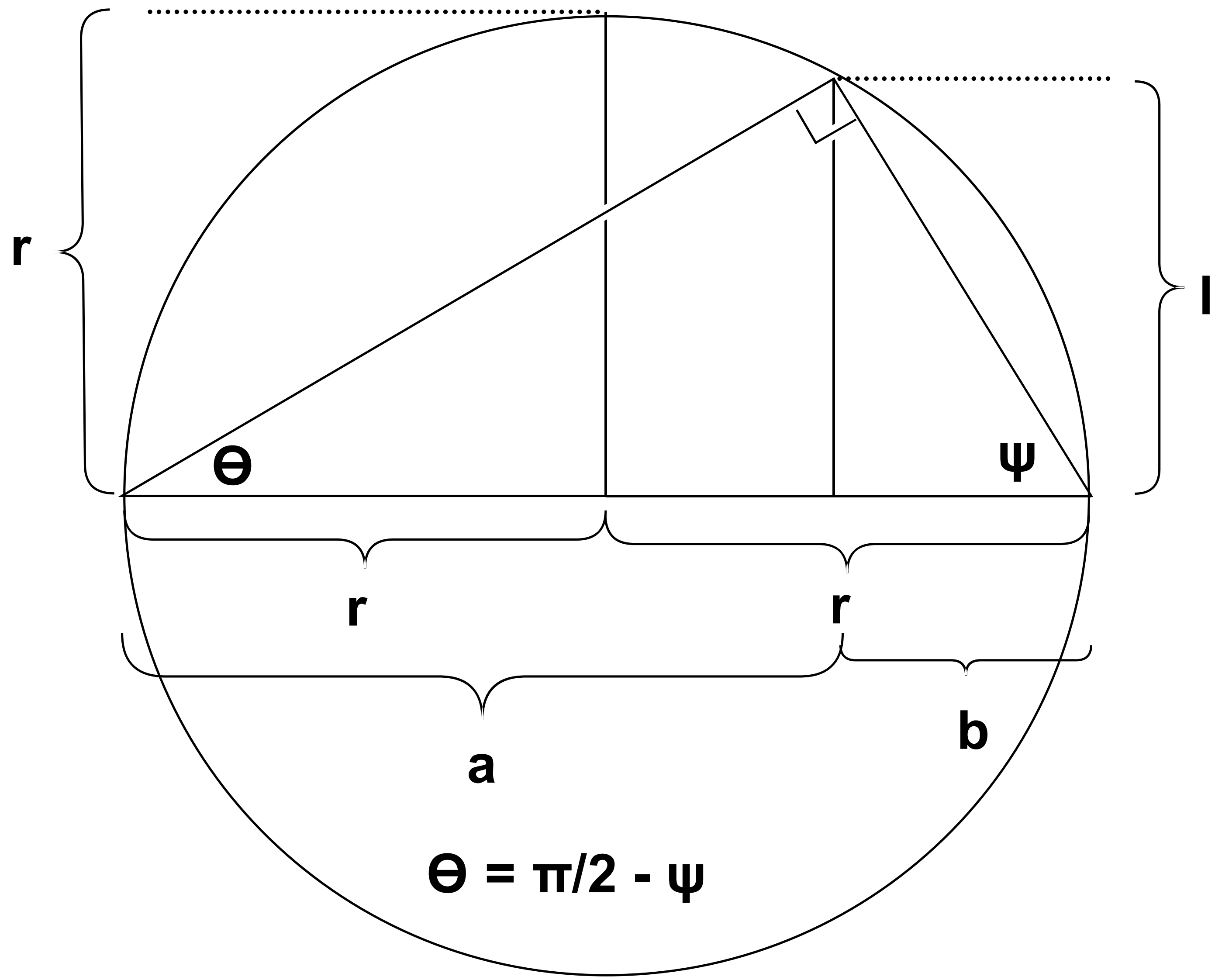
$$QM = AM \text{ only when all values of } x_i \text{ are identical}$$

Limited Proof that the Arithmetic Mean is greater than or equal to the Geometric Mean:

Limited form of the Thales Central Angle Theorem:



Limited proof that AM ≥ GM:



$$AM = \frac{a+b}{2} = r \qquad GM = \sqrt{ab}$$

$$\frac{l}{a} = tan(\theta) \qquad \frac{l}{b} = tan(\psi) \qquad \theta + \psi = \frac{\pi}{2} \qquad tan(\theta) = \cot(\psi)$$

$$\frac{l}{a} = \frac{b}{l} \qquad l^2 = ab \qquad l = \sqrt{ab} = GM$$

$$As\ l \leq r\ with\ equality\ only\ when\ a = b = r \qquad AM \geq GM$$

Full Proof that the Arithmetic Mean ≥ Geometric Mean:

$$Arithmetic\ Mean = AM = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \alpha$$

$$Geometric\ Mean = GM = (x_1 x_2 x_3 \cdots x_n)^{\frac{1}{n}}$$

With "α" being equal to the arithmetic mean, pick two values of x_i such that one is greater than α and one is less than α. Assume that x_1 and x_2 meet this criterion with x_2 being great than α and x_1 being less than α, as these can be assigned arbitrarily. If there is no value of x_i greater than α with at least one different value of x_i obligatorily less than α, then all the values of x_i are equal to α and "AM = GM". Substitute the values of "α" & "$x_1 + x_2 - \alpha$" into both the AM and GM:

$$AM = \frac{\alpha + (x_1 + x_2 - \alpha) + x_3 + \cdots + x_n}{n} = \alpha$$

$$GM = \{\alpha(x_2 + x_1 - \alpha)x_3 \cdots x_n\}^{\frac{1}{n}} = \{(\alpha x_2 + \alpha x_1 - \alpha^2)x_3 \cdots x_n\}^{\frac{1}{n}}$$

$$\alpha x_2 + \alpha x_1 - \alpha^2 - x_1 x_2 = (x_2 - \alpha)(\alpha - x_1)$$

$$x_2 - \alpha > 0 \quad \& \quad \alpha - x_1 > 0 \quad \& \quad \alpha x_2 + \alpha x_1 - \alpha^2 > x_1 x_2$$

The geometric mean has been increased by this substitution while the arithmetic mean has remained the same. Now reassign values:

$$Let\ x_1 + x_2 - \alpha = new\ value\ of\ x_1$$

$$Reassign\ x_3\ through\ x_n\ as\ x_2\ through\ x_{n-1}$$

$$New\ Arithmetic\ Mean = \frac{\alpha + x_1 + x_2 + x_3 + \cdots + x_{n-1}}{n} = \alpha$$

$$\alpha + x_1 + x_2 + x_3 + \cdots + x_{n-1} = n\alpha - \alpha = \alpha(n-1)$$

$$New\ Arithmetic\ Mean = \frac{x_1 + x_2 + x_3 + \cdots + x_{n-1}}{n-1} = \alpha$$

So now one can repeat the process. There must be at least one value of x_i greater than α and one value of x_i less than α unless all the values of x_1 through x_{n-1} are equal to α. We arbitrarily assign the value of x_2 to be greater than α and x_1 to be less than α. One continues the process over and over until all the values of x_i have been converted to α. Note that the final two values, assuming they are different from α, will both be converted to α.

$$New\ Arithmetic\ Mean = \frac{x_1 + x_2}{2} = \frac{(\alpha + z\alpha) + (\alpha - z\alpha)}{2} = \alpha$$

$$x_1 + x_2 - \alpha = \alpha + z\alpha + \alpha - z\alpha - \alpha = \alpha$$

With each step, the geometric mean increases while the arithmetic mean remains the same. Ultimately, the value of the geometric mean reaches α.

$$Increased\ Geometric\ Mean = (\alpha\alpha\alpha\alpha\alpha\alpha \cdots \alpha)^{\frac{1}{n}} = (\alpha^n)^{\frac{1}{n}} = \alpha$$

$$Original\ Arithmetic\ Mean = \alpha$$

The original geometric mean was thus less than or equal to the original arithmetic mean.

Proof that the Geometric Mean ≥ Harmonic Mean:

$$Geometric\ Mean = GM = (x_1 x_2 x_3 \cdots x_n)^{\frac{1}{n}}$$

$$Harmonic\ Mean = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \cdots + \frac{1}{x_n}}$$

Note that the harmonic mean is the inverse of the arithmetic mean of the inverses. The inverse of the geometric mean of the inverses is given below:

$$\frac{1}{Geometric\ Mean\ of\ inverses} = \frac{1}{\left(\left(\frac{1}{x_1}\right)\left(\frac{1}{x_2}\right)\left(\frac{1}{x_3}\right)\cdots\left(\frac{1}{x_n}\right)\right)^{\frac{1}{n}}}$$

$$\frac{1}{Geometric\ Mean\ of\ inverses} = \left(\frac{1}{\left(\left(\frac{1}{x_1}\right)\left(\frac{1}{x_2}\right)\left(\frac{1}{x_3}\right)\cdots\left(\frac{1}{x_n}\right)\right)}\right)^{\frac{1}{n}}$$

$$\frac{1}{Geometric\ Mean\ of\ inverses} = (x_1 x_2 x_3 \cdots x_n)^{\frac{1}{n}} = GM$$

By previous proofs:

$$\frac{1}{AM\ of\ inverses} \leq \frac{1}{GM\ of\ inverses} = (x_1 x_2 x_3 \cdots x_n)^{\frac{1}{n}} = GM$$

$$\frac{1}{AM\ of\ inverses} = HM \leq GM$$

The Weighted Average (Mean):

A weighted average (mean) is convenient when handling data for large populations. For instance, an average test score for a test graded on a scale of zero to one hundred (integers only) for 10,000 students will obviously contain many instances of students with the same grade (pigeon hole principle). Going forward, let "m" designate the value being averaged; in this case, the test grade. Rather than listing all the identical grades individually, one can use a weighted average.

$$n_i = number\ of\ data\ points\ with\ value\ m_i$$

$$\sum_{i=1}^{n} n_i = N = total\ number\ of\ data\ points$$

$$Weighted\ Arithmetic\ Mean = M_n = \frac{\sum_{i=1}^{n} n_i m_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i m_i}{N}$$

The weighted average can be restated with fractions:

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{N} = \sum_{i=1}^{n} x_i m_i \qquad x_i = \frac{n_i}{N}$$

Here is the weighted quadratic mean:

$$Weighted\ Quadratic\ Mean = Q_n = \sqrt{\frac{\sum_{i=1}^{n} n_i m_i^2}{\sum_{i=1}^{n} n_i}} = \sqrt{\frac{\sum_{i=1}^{n} n_i m_i^2}{N}}$$

$$Q_n = \sqrt{\frac{\sum_{i=1}^{n} n_i m_i^2}{N}} = \sum_{i=1}^{n} \sqrt{x_i m_i^2}$$

Here is the weighted geometric mean (does not simplify as much as the others):

$$G_n = \left\{\prod_{i=1}^{n} n_i m_i\right\}^{\frac{1}{n}}$$

Here is the weighted harmonic mean:

$$H_n = \frac{\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n} n_i \frac{1}{m_i}} = \frac{N}{\sum_{i=1}^{n} n_i \frac{1}{m_i}} = \frac{1}{\sum_{i=1}^{n} x_i \frac{1}{m_i}}$$

Note that the use of a weighted average does not change the QM ≥ AM ≥ GM ≥ HM relationship, as the weighting factor can be taken out and replaced with an individual listing of the combined data points to give the simpler form of the average. An example for illustration:

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{N} = \frac{3(5) + 4(6)}{7} = \frac{5+5+5+6+6+6+6}{7}$$

$$Q_n = \sqrt{\frac{\sum_{i=1}^{n} n_i m_i^2}{\sum_{i=1}^{n} n_i}} = \sqrt{\frac{2(3)^2 + 3(4)^2}{5}} = \sqrt{\frac{3^2 + 3^2 + 4^2 + 4^2 + 4^2}{5}}$$

Treating Data Points as Ratio Quantities:

One can keep track of units in all the different types of averages. Rather than averaging test scores as numbers, treat them as a ratio, *i.e.*, grade per person. If a student gets a grade of 80, then this is thought of as 80/person.

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} (\cancel{\#\ people})\frac{grade}{person}}{\cancel{total\ \#\ of\ people}} = units\ of\ \frac{grade}{person}$$

Some averaged quantities are normally thought of as ratios; such as a speed in miles/hour or price in dollars/pound.

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} (\cancel{\#\ pounds})\frac{dollars}{pound}}{\cancel{total\ \#\ of\ pounds}} = units\ of\ \frac{dollars}{pound}$$

Here is the unit analysis for the harmonic mean of test grades thought of as grade/person.

$$\frac{N}{\sum_{i=1}^{n} n_i \frac{1}{m_i}} = \frac{total\ \#\ of\ people}{(\#people)\frac{1}{\frac{grade}{person}}} = \frac{total\ \#\ of\ people}{(\#people)\frac{person}{grade}}$$

$$\frac{N}{\sum_{i=1}^{n} n_i \frac{1}{m_i}} = \frac{people}{\frac{(people)^2}{grade}} = \frac{grade}{person}$$

Note that a weighted average can be based on differing amounts of either the numerator quantity or denominator quantity of the ratio being averaged. One can determine an average price for differing numbers of pounds at differing prices in dollars per pound. This is a weighted average based on differing amounts of the denominator quantity (*e.g.*, pounds in dollars/pound). Conversely, a weighted average price can be obtained for differing amounts of dollars spent at differing prices; *i.e.*, an average based on differing amounts of the price's numerator quantity (*e.g.*, dollars in dollars/pound). For a weighted average of differing values present in differing amounts of the denominator quantity, use the weighted arithmetic mean:

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{\sum_{i=1}^{n} n_i} = \frac{(\#\ pounds)_i \left(\frac{dollars}{pound}\right)_i}{total\ \#\ of\ pounds} = \left(\frac{dollars}{pound}\right)_{average}$$

For a weighted average based on different amounts of the numerator quantity, use the weighted harmonic mean:

$$H_n = \frac{\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n} n_i \frac{1}{m_i}} = \frac{total\ \$\ spent}{(\$\ at\ price)_i \frac{1}{\left(\frac{\$\ price}{pound}\right)_i}}$$

$$H_n = \frac{total\ \$}{(\$)_i \left(\frac{pounds}{\$}\right)_i} = \left(\frac{\$}{pound}\right)_{average}$$

<u>When to use the different types of averages:</u>

<u>The quadratic mean</u> is used when both positive and negative numbers must be considered based on their magnitude. The standard deviation for a population is a quadratic mean of deviations from the arithmetic mean. The quadratic mean is an important concept when considering alternating current, though it usually applied continuously with calculus.

<u>The arithmetic mean</u> yields the central value of a finite arithmetic series (*i.e.*, mean = median). We will designate the arithmetic mean (AM) as M_n for the remainder of this paper. The arithmetic mean of the finite sequence 3, 6, 9, 12, 15 is equal to 9 ($M_n = 9$). If there are an even number of terms in a finite arithmetic sequence, then the arithmetic mean is equal to the arithmetic mean of the central two terms. For instance, the arithmetic mean of the finite sequence 3, 6, 9, 12 is equal to the arithmetic mean of 6 & 9 (*i.e.*, 7.5). When the arithmetic mean (M_n) is multiplied by the total number of data points (N), one gets the total sum of all the data points.

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i m_i}{N} \qquad NM_n = \sum_{i=1}^{n} n_i m_i$$

<u>The geometric mean</u> yields the central term of a finite geometric series (*i.e.*, mean = median). For instance, the geometric mean of the finite geometric sequence 3, 9, 27 81, 243 is equal to 27. If there are an even number of terms in a finite geometric series, then the geometric mean is equal to the geometric mean of the central two terms. For instance, the geometric mean of the finite sequence 3, 9, 27, 81 is equal to the geometric men of 9 & 27 (*i.e.*, $\sqrt{(9)(27)} = \sqrt{243} \approx 15.6$). The geometric mean should be used for data that accrues geometrically, such as the average annual change of an investment fund. Negative numbers can't be used in geometric means. The value of terms to be averaged geometrically should all be positive. One must convert a series containing negative terms to all positive values, possibly by converting the terms from values representing change to values representing remaining amount. For instance, an investment that gains 30% in year one, loses 20% in year two, and finally gains 15% in year three should be collated as 1.3, 0.80, 1.15 and not as 30, -20, 15. The geometric mean of 1.3, 0.80, 1.15 equals $\{(1.3)(0.80)(1.15)\}^{\frac{1}{3}} = (1.196)^{\frac{1}{3}} \approx 1.0615$ (average annual interest rate of approximately 6.15% yielding 19.6% gain after three years).

The harmonic mean yields the "full mediant" of a finite sequence of values wherein the numerators of all the values in the series are made identical. The full mediant of a finite sequence of terms is obtained by dividing the sum of the all the numerators by the sum of the all the denominators. Consider:

$$\frac{2}{3}, \frac{2}{5}, \frac{2}{9}, \frac{2}{10}, \frac{2}{11}, \frac{2}{11}$$

$$HM = \frac{6}{\frac{3}{2} + \frac{5}{2} + \frac{9}{2} + \frac{10}{2} + \frac{11}{2} + \frac{11}{2}} = \frac{6(2)}{3 + 5 + 9 + 10 + 11 + 11}$$

Here is another example:

$$3,5,7,\frac{2}{3},11 = \frac{2}{\frac{2}{3}}, \frac{2}{\frac{2}{5}}, \frac{2}{\frac{2}{7}}, \frac{2}{3}, \frac{2}{\frac{2}{11}}$$

$$HM = \frac{5}{\frac{2/3}{2} + \frac{2/5}{2} + \frac{2/7}{2} + \frac{3}{2} + \frac{2/11}{2}} = \frac{5(2)}{\frac{2}{3} + \frac{2}{5} + \frac{2}{7} + 3 + \frac{2}{11}}$$

Note that when the denominators form an arithmetic series, assuming a finite sequence of values wherein all the numerators are equal, then the harmonic mean is equal to the central term in the series (*i.e.*, mean = median) or, in the case where there are an even number of terms, the mediant (equivalent to the harmonic mean) of the central two terms.

$$\frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \frac{1}{9}, \frac{1}{11}$$

$$HM = \frac{5}{3 + 5 + 7 + 9 + 11} = \frac{1}{7}$$

$$\frac{3}{5}, \frac{3}{8}, \frac{3}{11}, \frac{3}{14}$$

$$HM = \frac{4}{\frac{5}{3} + \frac{8}{3} + \frac{11}{3} + \frac{14}{3}} = \frac{4(3)}{5 + 8 + 11 + 14} = \frac{12}{38} = \frac{6}{19} = \frac{3 + 3}{8 + 11}$$

As previously shown, the weighted harmonic mean is used when one is calculating a weighted average with weighting by differing amounts of the numerator quantity of the ratio values being averaged. Note that using a weighted harmonic mean in this way provides a value which is identical to the arithmetic mean that would have been obtained had the weighting been done by equivalent amounts (*i.e.*, equivalent to the amount, not the same value) of the denominator quantity of the ratio value being averaged.

A speed in miles per hour is an example of a ratio value. If one wants a weighted average for different speeds (miles per hour) traveled for differing amounts of time (hours – the denominator of the ratio value being averaged), then one uses a weighted arithmetic mean.

$$t_i = time\ at\ speed\ i \quad s_i = speed\ i \quad T = total\ time = \sum_{i=1}^{n} t_i$$

$$M_n = \frac{\sum_{i=1}^{n} t_i s_i}{\sum_{i=1}^{n} t_i} = \sum_{i=1}^{n} \frac{t_i s_i}{T}$$

The harmonic mean is used when the ratio value is weighted by variable amounts of the ratio's numerator quantity (distance in miles for speed in miles/hour).

$$l_i = distance\ at\ speed\ i \quad s_i = speed\ i \quad L = total\ distance = \sum_{i=1}^{n} l_i$$

$$H_n = \frac{\sum_{i=1}^{n} l_i}{\sum_{i=1}^{n} l_i \frac{1}{s_i}} = \frac{L}{\sum_{i=1}^{n} l_i \frac{1}{s_i}}$$

Note that this harmonic mean is equal to the arithmetic mean that would be obtained if the weighting was in times (hours) corresponding to the distances (miles).

If the weightings based on a denominator quantity are all equal, then the weightings can be replaced as division by n (number of data points) yielding the simple arithmetic mean.

$$M_n = \frac{\sum_{i=1}^{n} t_i s_i}{\sum_{i=1}^{n} t_i} = \frac{t_i}{n t_i}\sum_{i=1}^{n} s_i = \frac{1}{n}\sum_{i=1}^{n} s_i$$

If the weightings based on a numerator quantity are all equal, then the weightings can be replaced by a numerator equal to “n” (number of data points) yielding the simple harmonic mean.

$$H_n = \frac{\sum_{i=1}^{n} l_i}{\sum_{i=1}^{n} l_i \frac{1}{s_i}} = \frac{n l_i}{l_i \sum_{i=1}^{n} \frac{1}{s_i}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{s_i}}$$

Measures of Dispersity:

The variance and the standard deviation are two common measures of the dispersity of a data set.

$$s = \sqrt{\frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{\sum_{i=1}^{n} n_i}} = \sqrt{\frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{N}} = stand. dev.$$

$$s^2 = \frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{N} = variance$$

Another measure of dispersity known as the “Polydispersity Index” is used in polymer science. When a polymer is produced, there will be a range of molecular weights (MW’s) produced. The number average molecular weight of the polymer can be calculated as:

$$n_i = number\ of\ molecules\ (moles) with\ GMW\ of\ m_i$$

$$N = total\ number\ of\ molecules\ (moles)$$

$$x_{i,n} = number\ (mole)\ fraction\ of\ fraction = \frac{n_i}{N}$$

$$M_n = \frac{\sum_{i=1}^{n} n_i m_i}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i m_i}{N} = \sum_{i=1}^{n} x_{i,n} m_i$$

The number average molecular weight is a traditional weighted arithmetic mean. One can also calculate a weight average molecular weight as M_w:

$$NM_n = \sum_{i=1}^{n} n_i m_i = total\ weight\ of\ polymer$$

$$M_w = \frac{\frac{\sum_{i=1}^{n} n_i m_i^2}{N}}{M_n} = \frac{\sum_{i=1}^{n} n_i m_i^2}{NM_n} = \frac{\frac{\sum_{i=1}^{n} n_i m_i^2}{N}}{\frac{\sum_{i=1}^{n} n_i m_i}{N}} = \frac{\sum_{i=1}^{n} n_i m_i^2}{\sum_{i=1}^{n} n_i m_i} = \frac{\sum_{i=1}^{n} x_i m_i^2}{\sum_{i=1}^{n} x_i m_i}$$

$$M_w = \frac{\sum_{i=1}^{n} n_i m_i m_i}{NM_n} = \frac{\sum_{i=1}^{n} (weight\ of\ fraction) m_i}{(total\ weight\ of\ polymer)}$$

$$M_w = \sum_{i=1}^{n} x_{i,w} m_i \qquad x_{i,w} = weight\ fraction$$

The weight average molecular weight is an average of the molecular weights weighted by their weight fraction rather than the number fraction. Note that this type of “extended” average can be calculated for any ratio quantity. Consider a similar set of weighted averages for a set of apple prices in dollars per pound for differing weights of apples.

$$p_i = price\ of\ apples\ for\ fraction\ "i"\ in\ dollars/pound$$

$$w_i = weight\ of\ apples\ for\ fraction\ i\ in\ pounds$$

$$w_i p_i = dollars\ spent\ to\ buy\ w_i\ pounds\ at\ price\ p_i$$

$$\sum_{i=1}^{n} w_i p_i = total\ dollars\ spent = WM_n$$

$$\sum_{i=1}^{n} w_i = total\ weight\ of\ apples = W$$

$$M_n = \frac{\sum_{i=1}^{n} w_i p_i}{\sum_{i=1}^{n} w_i} = \frac{\sum_{i=1}^{n} w_i p_i}{W} = \sum_{i=1}^{n} x_{i,n} p_i = \frac{total\ dollars\ spent}{total\ weight}$$

$$x_{i,n} = weight\ fraction\ of\ fraction$$

The weight average for this type of data set would be:

$$M_w = \frac{\sum_{i=1}^{n} w_i p_i p_i}{W M_n} = \frac{\sum_{i=1}^{n} (price\ of\ fraction) p_i}{(total\ price)} = \sum_{i=1}^{n} x_{i,w} p_i$$

$$x_{i,w} = price\ fraction\ of\ fraction$$

The weight average apple price is an average of the apple prices weighted by their price fraction rather than their weight fraction.

Relationship of Mn & Mw to the variance:

$$s^2 = \frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{N} = variance$$

$$\frac{\sum_{i=1}^{n} n_i (M_n - m_i)^2}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i (M_n^2 - 2M_n m_i + m_i^2)}{N}$$

$$s^2 = \frac{\sum_{i=1}^{n} n_i M_n^2}{N} - 2\frac{\sum_{i=1}^{n} n_i m_i M_n}{N} + \frac{\sum_{i=1}^{n} n_i m_i^2}{N}$$

$$\frac{s^2}{M_n^2} = \frac{\sum_{i=1}^{n} n_i M_n^2}{N M_n^2} - 2\frac{\sum_{i=1}^{n} n_i m_i M_n}{N M_n^2} + \frac{\sum_{i=1}^{n} n_i m_i^2}{N M_n^2}$$

$$\left(\frac{s}{M_n}\right)^2 = \frac{\sum_{i=1}^{n} n_i}{N} - 2\frac{\sum_{i=1}^{n} n_i m_i}{N M_n} + \frac{\sum_{i=1}^{n} n_i m_i^2}{N M_n M_n}$$

$$\left(\frac{s}{M_n}\right)^2 = 1 - 2 + \frac{M_w}{M_n} = \frac{M_w}{M_n} - 1$$

$$\left(\frac{s}{M_n}\right)^2 = \frac{M_w}{M_n} - 1$$

$$\frac{M_w}{M_n} = polydispersity\ index$$