

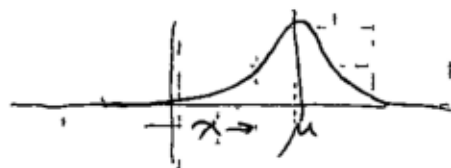
My notes on some math subjects (some sections include book Xerox pages and web pages)

1. Standard Distributions Section
Gaussian page, Binomial/Poisson notes + Proakis notes + Proakis Xerox
(Proakis has Binomial, Gaussian, Chi-square, Rayleigh, Rice) + self notes
2. Basic Probability theory (web) + Balls in Bins
- ✓ 3. The Levi-Civita Symbol
4. The Method of Stationary Phase (for evaluating integrals) + web
5. Selected integrals
Remainder Function, Mod etc
Gauss's law like integrals in θ, ϕ space.
6. Misc Single page items [Int, Rem; $\int d\Omega (\vec{\nabla} \cdot \vec{f}) = 0$; Hilbert Space; multi integrals;]
7. Asymptotic Series (web)
8. Gaussian Quadrature
9. Mean Value Theorems + web
10. The Simplex Method (Linear Programming Theory)
11. Rubber Sheet described by Laplace.
12. Laplace Method for ODE's
13. Non-Linear ODE's (Olver)
- 14.
- 15.
- 16.
- 17.

Normal (Gaussian) Distribution

8-24-03

$$g(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$$



The moments are:

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx x^n g(x, \mu, \sigma^2)$$

The first several are:

$$\langle x^0 \rangle = 1$$

$$\langle x^1 \rangle = \langle x \rangle = \mu$$

$$\langle x^2 \rangle = \langle x^2 \rangle = \sigma^2 + \mu^2 = \text{variance from } \phi$$

$$\langle x^3 \rangle = \langle x^3 \rangle = \mu(\mu^2 + 3\sigma^2)$$

$$\langle x^4 \rangle = \langle x^4 \rangle = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

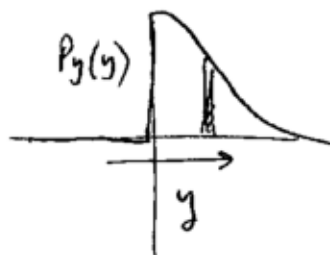
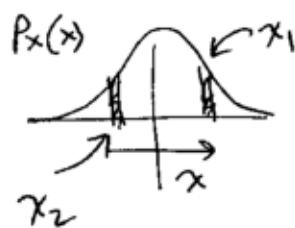
Now consider

$$\langle x^2 \rangle - \langle x \rangle^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 = \text{variance from mean.}$$

$$\langle x^4 \rangle - \langle x^2 \rangle^2 = 2\sigma^2(\sigma^2 + 2\mu^2)$$

Example #1

"Rician + all that..."



$$y = x^2$$

$$x_1 = +\sqrt{y}$$

$$x_2 = -\sqrt{y}$$

$$\frac{dx_1}{dy} = \frac{1}{2\sqrt{y}}$$

$$\frac{dy}{dx} = 2x$$

Derivation

$$P_Y(y)|dy| = P_X(x_1)|dx_1| + P_X(x_2)|dx_2|$$

$$\Rightarrow P_Y(y) = P_X(x_1)\left|\frac{dx_1}{dy}\right| + P_X(x_2)\left|\frac{dx_2}{dy}\right|$$

$$= P_X(\sqrt{y})\frac{1}{2\sqrt{y}} + P_X(-\sqrt{y})\frac{1}{2\sqrt{y}}$$

Rewrite as:

$$P_Y(y) = \frac{P_X(x_1)}{\left|\frac{dy}{dx}\right|_{x=x_1}} + \frac{P_X(x_2)}{\left|\frac{dy}{dx}\right|_{x=x_2}} = \frac{P_X(+\sqrt{y})}{|2x_1|} + \frac{P_X(-\sqrt{y})}{|2x_2|} \checkmark$$

General result:

$$y = g(x) = y(x)$$

There are n real roots: $x_1(y), x_2(y), \dots, x_n(y)$.

Then

$$P_Y(y) = \sum_{i=1}^n \frac{P_X(x_i)}{\left|\frac{dy}{dx}\right|_{x=x_i}} \quad (2-1-49)$$

Example #2

Based on Example #1. Suppose $y = x^2$ and

$$P_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)} = \text{gaussian}$$

Then

$$\begin{aligned} P_Y(y) &= \frac{P_X(\sqrt{y}) + P_X(-\sqrt{y})}{2\sqrt{y}} = \frac{P_X(\sqrt{y})}{\sqrt{y}} \\ &= \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{\sqrt{y}} e^{-y/(2\sigma^2)} = \frac{1}{\sqrt{2\pi y}\sigma} e^{-y/(2\sigma^2)} \\ &= \text{as shown in (2-1-105)}. \quad \checkmark \end{aligned}$$

Example #3

$$R = \sqrt{\sum_{i=1}^n X_i^2}$$

If X_i has $p(x) = \text{gaussian } \mu=0, \sigma^2$, then

$$P_R(r) = \frac{r^{n-1}}{2^{\frac{n-2}{2}} \sigma^n \Gamma(\frac{n}{2})} e^{-r^2/(2\sigma^2)} \quad (2-1-139)$$

$$\langle r^k \rangle = (2\sigma^2)^{\frac{k}{2}} \frac{\Gamma(\frac{n+k}{2})}{\Gamma(\frac{n}{2})} \quad (2-1-139) \rightarrow \left(\frac{n}{2}\right)^{k/2} \text{ as } n \rightarrow \infty.$$

AS p 257 6.1.39

So for example

$$k=1 \quad \checkmark \langle r \rangle = (2\sigma^2)^{\frac{1}{2}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \equiv d_n \quad \lim_{n \rightarrow \infty} d_n = \sqrt{\frac{n}{2}} \quad \begin{matrix} \text{A+S} \\ \text{p 257} \\ \text{6.1.39} \\ \text{+ Maple} \end{matrix}$$

$$k=2 \quad \checkmark \langle r^2 \rangle = (2\sigma^2) \cdot \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n}{2})} = 2\sigma^2 \cdot \frac{n}{2} = n\sigma^2 \quad \checkmark$$

Then:

$$\langle r^2 \rangle - \langle r \rangle^2 = (n\sigma^2) - (2\sigma^2) d_n^2 = n\sigma^2 - 2\sigma^2 d_n^2 = \sigma^2 [n - 2 d_n^2]$$

And

$$\sqrt{\langle r^2 \rangle - \langle r \rangle^2} = \sqrt{n\sigma^2 - 2d_n^2\sigma^2} = \text{STD}(R)$$

$$\tilde{R} = 2 \sqrt{\sum x_i^2} \rightarrow \sqrt{\frac{1}{n}} \cdot \sqrt{\sum x_i^2} = \sqrt{\frac{1}{n}}$$

$$\tilde{P}_R(r) = \frac{1}{2} P_R\left(\frac{r}{2}\right)$$

$$\langle \tilde{r}^k \rangle = \int_{-\infty}^{\infty} dr r^k \tilde{P}_R(r) = \int_{-\infty}^{\infty} dr r^k \frac{1}{2} P_R\left(\frac{r}{2}\right).$$

$$\text{Let } r' = r/2 \quad dr' = \frac{1}{2} dr$$

$$= \int_{-\infty}^{\infty} 2dr' (2r')^k \frac{1}{2} P_R(r')$$

$$= 2^k \langle r^k \rangle.$$

$$\widetilde{\langle r^2 \rangle - \langle r \rangle^2} = 2^2 [n^2 \sigma^4 - 2\sigma^2 \alpha n^2].$$

So if $2 \Rightarrow \frac{1}{\sqrt{n}}$. Then.

$$\langle \tilde{r}^k \rangle = \left(\frac{1}{\sqrt{n}}\right)^k \langle r^k \rangle$$

$$\widetilde{\langle r^2 \rangle - \langle r \rangle^2} = \left(\frac{1}{\sqrt{n}}\right)^2 [n^2 \sigma^4 - 2\sigma^2 \alpha n^2].$$

$$= \frac{1}{n} [n^2 \sigma^4 - 2\sigma^2 \frac{n}{2}]$$

$$= e^{\left[\ln P\left(\frac{n+1}{2}\right) - \ln P\left(\frac{n}{2}\right) \right]} \quad \frac{P\left(\frac{n+1}{2}\right)}{P\left(\frac{n}{2}\right)} = \frac{e^{\ln P\left(\frac{n+1}{2}\right)}}{e^{\ln P\left(\frac{n}{2}\right)}}$$

cdf = cumulative dist. fcn
(area so far)
pdf = prob. density fcn

This conditional cdf satisfies the properties previously established for these functions, such as

$$F(x_1, x_2, \dots, x_k | x_{k+1}, \dots, x_n) = F(x_2, x_3, \dots, x_k | x_{k+1}, \dots, x_n)$$

$$F(-\infty, x_2, \dots, x_k | x_{k+1}, \dots, x_n) = 0$$

Statistically Independent Random Variables. We have already defined statistical independence of two or more events of a sample space S . The concept of statistical independence can be extended to random variables defined on a sample space generated by a combined experiment or by repeated trials of a single experiment. If the experiments result in mutually exclusive outcomes, the probability of an outcome in one experiment is independent of an outcome in any other experiment. That is, the joint probability of the outcomes factors into a product of the probabilities corresponding to each outcome. Consequently, the random variables corresponding to the outcomes in these experiments are independent in the sense that their joint pdf factors into a product of marginal pdfs. Hence the multidimensional random variables are statistically independent if and only if

$$F(x_1, x_2, \dots, x_n) = F(x_1)F(x_2) \cdots F(x_n) \quad (2-1-37)$$

or, alternatively,

$$p(x_1, x_2, \dots, x_n) = p(x_1)p(x_2) \cdots p(x_n) \quad (2-1-38)$$

2-1-2 Functions of Random Variables

A problem that arises frequently in practical applications of probability is the following. Given a random variable X , which is characterized by its pdf $p(x)$, determine the pdf of the random variable $Y = g(X)$, where $g(X)$ is some given function of X . When the mapping g from X to Y is one-to-one, the determination of $p(y)$ is relatively straightforward. However, when the mapping is not one-to-one, as is the case, for example, when $Y = X^2$, we must be very careful in our derivation of $p(y)$.

Example 2-1-1 ✓

Consider the random variable Y defined as

$$Y = aX + b \quad (2-1-39)$$

where a and b are constants. We assume that $a > 0$. If $a < 0$, the approach is similar (see Problem 2-3). We note that this mapping, illustrated in Fig. 2-1-4(a) is linear and monotonic. Let $F_X(x)$ and $F_Y(y)$ denote the cdfs for X and Y , respectively.† Then

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) \checkmark$$

$$= \int_{-\infty}^{(y-b)/a} p_X(x) dx = F_X\left(\frac{y-b}{a}\right) \checkmark \quad (2-1-40)$$

† To avoid confusion in changing variables, subscripts are used in the respective pdfs and cdfs.

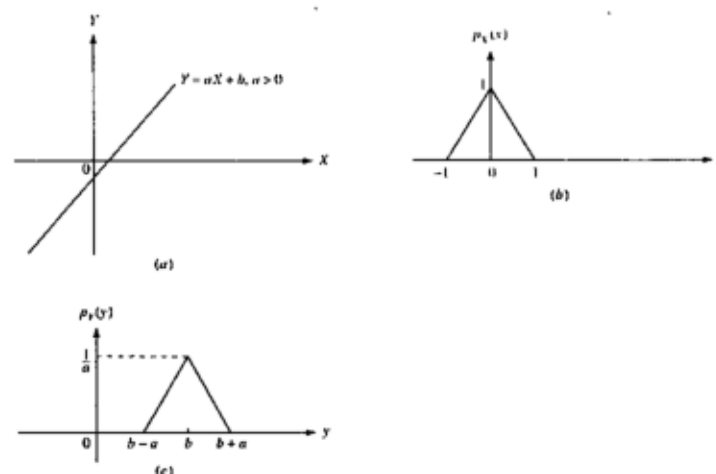


FIGURE 2-1-4 A linear transformation of a random variable X and an example of the corresponding pdfs of X and Y .

By differentiating (2-1-40) with respect to y , we obtain the relationship between the respective pdfs. It is

$$p_Y(y) = \frac{1}{a} p_X\left(\frac{y-b}{a}\right) \checkmark \quad (2-1-41)$$

Thus (2-1-40) and (2-1-41) specify the cdf and pdf of the random variable Y in terms of the cdf and pdf of the random variable X for the linear transformation in (2-1-39). To illustrate this mapping for a specific pdf $p_X(x)$, consider the one shown in Fig. 2-1-4(b). The pdf $p_Y(y)$ that results from the mapping in (2-1-39) is shown in Fig. 2-1-4(c).

Example 2-1-2

Consider the random variable Y defined as

$$Y = aX^3 + b, \quad a > 0 \quad (2-1-42)$$

As in Example 2-1-1, the mapping between X and Y is one-to-one. Hence

$$F_Y(y) = P(Y \leq y) = P(aX^3 + b \leq y) \checkmark$$

$$= P\left[X \leq \left(\frac{y-b}{a}\right)^{1/3}\right] = F_X\left[\left(\frac{y-b}{a}\right)^{1/3}\right] \checkmark \quad (2-1-43)$$

WOLFRAM RESEARCH **mathworld.wolfram.com** OTHER WOLFRAM

ASTRONOMY BIOGRAPHY CHEMISTRY MATHEMATICS PHYSICS

Search Site Go

ERIC WEIGSTEIN'S
world of
MATHEMATICS

[Probability and Statistics](#) ▶ [Statistical Distributions](#) ▶ [Discrete Distributions](#) ▶
[Math Contributors](#) ▶ [Trott](#) ▼

Binomial Distribution

INDEX BY SUBJECT

- ◊ Algebra
- ◊ Applied Mathematics
- ◊ Calculus and Analysis
- ◊ Discrete Mathematics
- ◊ Foundations of Mathematics
- ◊ Geometry
- ◊ History and Terminology
- ◊ Number Theory
- ◊ Probability and Statistics
- ◊ Recreational Mathematics
- ◊ Topology

ALPHABETICAL INDEX ◊

◊ ABOUT THIS SITE

◊ AUTHOR'S NOTE

◊ FAQ

◊ WHAT'S NEW

◊ RANDOM ENTRY

◊ BE A CONTRIBUTOR

◊ SIGN THE GUESTBOOK

◊ EMAIL COMMENTS

◊ HOW CAN I HELP?

◊ TERMS OF USE

ERIC'S OTHER SITES ◊

◊ ORDER BOOK FROM
AMAZON.COM

DOWNLOAD

Mathematica Notebook MathReader



The binomial distribution gives the discrete probability distribution $P_p(n|N)$ of obtaining exactly n successes out of N Bernoulli trials (where the result of each Bernoulli trial is true with probability p and false with probability $q = 1 - p$). The binomial distribution is therefore given by

$$P_p(n|N) = \binom{N}{n} p^n q^{N-n} \quad (1)$$

$$= \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}, \quad (2)$$

where $\binom{N}{n}$ is a binomial coefficient. The above plot shows the distribution of n successes out of $N = 20$ trials with $p = q = 1/2$.

The probability of obtaining *more* successes than the n observed in a binomial distribution is

$$P = \sum_{k=n+1}^N \binom{N}{k} p^k (1-p)^{N-k} = I_p(n+1, N-n), \quad (3)$$

where

$$I_x(a, b) \equiv \frac{B(x; a, b)}{B(a, b)}, \quad (4)$$

Chapter 7

Probability

7.1 Experiments, Sample Spaces and Events

Start with some definitions we will need in our study of probability.

An EXPERIMENT is an activity with an observable result. Tossing coins, rolling dice and choosing cards are all probability experiments.

The result of the experiment is called the **OUTCOME** or **SAMPLE POINT**. So the two possible outcomes from tossing a coin are H (heads) and T (tails).

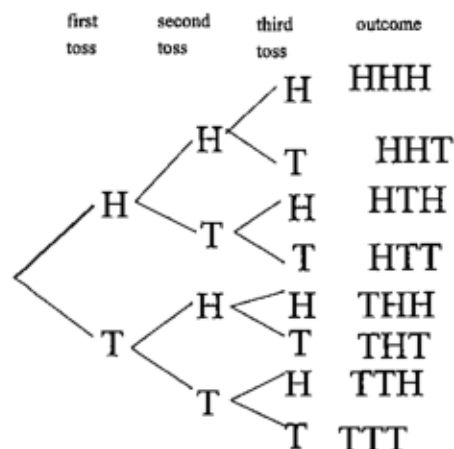
The set of all outcomes or sample points is called the SAMPLE SPACE of the experiment.

An EVENT is a subset of a sample space. That is, an event can contain one or more outcomes that are in the sample space.

Consider tossing a coin. The sample space is $S = \{H, T\}$. The events that are possible in this experiment are $\emptyset$, $\{H\}$, $\{T\}$, S . So, while there are 2 outcomes in the sample space, there are 4 different events. ✓

If a 6-sided die is rolled, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$.

Sometimes we use a tree diagram to find all the possible outcomes of an experiment. Consider tossing a coin 3 times and noting the result of each toss.



So we find there are 8 elements in the sample space. We would show this as

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \checkmark$$

Levi-Civita symbol and cross product vector/tensor

Patrick Guio

\$Id: levi-civita.tex,v 1.9 2001/08/29 15:53:02 patricg Exp \$

Definitions

The Levi-Civita symbol ϵ_{ijk} is a tensor of rank three and is defined by

$$\epsilon_{ijk} = \begin{cases} 0, & \text{if any two labels are the same} \\ 1, & \text{if } i, j, k \text{ is an even permutation of } 1, 2, 3 \\ -1, & \text{if } i, j, k \text{ is an odd permutation of } 1, 2, 3 \end{cases} \quad (1)$$

The Levi-Civita symbol ϵ_{ijk} is anti-symmetric on each pair of indexes.

The determinant of a matrix A with elements a_{ij} can be written in term of ϵ_{ijk} as

$$\det \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \epsilon_{ijk} a_{1i} a_{2j} a_{3k} = \epsilon_{ijk} a_{1i} a_{2j} a_{3k} \quad \leftarrow \quad (2)$$

Note the compact notation where the summation over the spatial directions is dropped. It is this one that is in use.

Note that the Levi-Civita symbol can therefore be expressed as the determinant, or mixed triple product, of any of the unit vectors $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ of a normalised and direct orthogonal frame of reference.

$$\epsilon_{ijk} = \det(\hat{e}_i, \hat{e}_j, \hat{e}_k) = \hat{e}_i \cdot (\hat{e}_j \times \hat{e}_k) \quad (3)$$

Now we can define by analogy to the definition of the determinant an additional type of product, the vector product or simply cross product

$$\mathbf{a} \times \mathbf{b} = \det \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \epsilon_{ijk} \hat{e}_i a_j b_k \quad (4)$$

for each coordinate

$$(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k \quad (5)$$

Properties

- The Levi-Civita tensor ϵ_{ijk} has $3 \times 3 \times 3 = 27$ components.
- $3 \times (6 + 1) = 21$ components are equal to 0.
- 3 components are equal to 1.
- 3 components are equal to -1.

Identities

The product of two Levi-Civita symbols can be expressed as a function of the Kronecker's symbol δ_{ij}

$$\begin{aligned}\epsilon_{ijk}\epsilon_{lmn} = & +\delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \\ & -\delta_{in}\delta_{jl}\delta_{kn} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl}\end{aligned}\quad (6)$$

Setting $i = l$ gives

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\quad (7)$$

Proof

$$\begin{aligned}\epsilon_{ijk}\epsilon_{imn} &= \delta_{ii}(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) + \delta_{im}\delta_{jn}\delta_{ki} + \delta_{in}\delta_{ji}\delta_{km} - \delta_{im}\delta_{ji}\delta_{kn} - \delta_{in}\delta_{jm}\delta_{ki} \\ &= 3(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) + \delta_{km}\delta_{jn} + \delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn} - \delta_{kn}\delta_{jm} \\ &= \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\end{aligned}$$

Setting $i = l$ and $j = m$ gives

$$\epsilon_{ijk}\epsilon_{ijn} = 2\delta_{kn}\quad (8)$$

Setting $i = l$, $j = m$ and $k = n$ gives

$$\epsilon_{ijk}\epsilon_{ijk} = 6\quad (9)$$

Therefore

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})\quad (10)$$

proof

Physics 209

"Jackson" class.

Fall 2002

Notes 3

The Levi-Civita Symbol

The Levi-Civita symbol is useful for converting cross products and curls into the language of tensor analysis, and for many other purposes. The following is a summary of its most useful properties in three-dimensional Euclidean space.

The Levi-Civita symbol is defined by

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } (ijk) \text{ is an even permutation of } (123); \\ -1, & \text{if } (ijk) \text{ is an odd permutation of } (123); \\ 0, & \text{otherwise.} \end{cases} \quad (3.1)$$

It has 27 components, of which only 6 are nonzero. It follows directly from this definition that ϵ_{ijk} changes sign if any two of its indices are exchanged,

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji}. \quad (3.2)$$

The Levi-Civita symbol is convenient for expressing cross products and curls in tensor notation. For example, if $\mathbf{A}$ and $\mathbf{B}$ are two vectors, then

$$(\mathbf{A} \times \mathbf{B})_i = \epsilon_{ijk} A_j B_k, \quad (3.3)$$

and

$$(\nabla \times \mathbf{B})_i = \epsilon_{ijk} \frac{\partial B_k}{\partial x_j}. \quad (3.4)$$

Any combination of an even number of Levi-Civita symbols (or an even number of cross products and curls) can be reduced to dot products with the following system of identities. Similarly, any combination of an odd number of Levi-Civita symbols (or an odd number of cross products and curls) can be reduced to a single Levi-Civita symbol (or a cross product or a curl) plus dot products. The first is the most general:

$$\epsilon_{ijk} \epsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}. \quad (3.5)$$

Notice that the indices (ijk) label the rows, while (lmn) label the columns. If this is contracted on i and l , we obtain

$$\epsilon_{ijk} \epsilon_{imn} = \begin{vmatrix} \delta_{jm} & \delta_{jn} \\ \delta_{km} & \delta_{kn} \end{vmatrix} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}. \quad (3.6)$$

This identity is the one used most often, for boiling down two cross products that have one index in common, such as $\nabla \times (\mathbf{A} \times \mathbf{B})$. By contracting Eq. (3.6) in j and m we obtain

$$\epsilon_{ijk} \epsilon_{ijn} = 2\delta_{kn}. \quad (3.7)$$

Finally, contracting on k and n we obtain

$$\epsilon_{ijk} \epsilon_{ijk} = 6. \quad (3.8)$$

It should be clear how to generalize these identities to higher dimensions.

If $A_{ij} = -A_{ji}$ is an antisymmetric, 3×3 tensor, it has 3 independent components that we can associate with a 3-vector $\mathbf{A}$, as follows:

$$A_{ij} = \begin{pmatrix} 0 & A_3 & -A_2 \\ -A_3 & 0 & A_1 \\ A_2 & -A_1 & 0 \end{pmatrix} = \epsilon_{ijk} A_k. \quad (3.9)$$

The inverse of this is

$$A_{ij} = \frac{1}{2} \epsilon_{ijk} A_k. \quad (3.10)$$

Using these identities, the multiplication of an antisymmetric matrix times a vector can be reexpressed in terms of a cross product. That is, if

$$X_i = A_{ij} Y_j \quad (3.11)$$

then

$$\mathbf{X} = \mathbf{Y} \times \mathbf{A}. \quad (3.12)$$

Similarly, if $\mathbf{A}$ and $\mathbf{B}$ are two vectors, then

$$A_i B_j - A_j B_i = \epsilon_{ijk} (\mathbf{A} \times \mathbf{B})_k, \quad (3.13)$$

and

$$\frac{\partial B_j}{\partial x_i} - \frac{\partial B_i}{\partial x_j} = \epsilon_{ijk} (\nabla \times \mathbf{B})_k. \quad (3.14)$$

Finally, if M_{ij} is a 3×3 matrix (or tensor), then

$$\det M = \epsilon_{ijk} M_{1i} M_{2j} M_{3k} = \frac{1}{6} \epsilon_{ijk} \epsilon_{lmn} M_{il} M_{jm} M_{kn}. \quad (3.15)$$

The Levi-Civita symbol has been defined here only on $\mathbb{R}^3$, but most of the properties above are easily generalized to $\mathbb{R}^n$ (including the case $n = 2$). It only transforms as a tensor under proper orthogonal changes of coordinates, which is why we are calling it a “symbol” instead of a “tensor.” It can, however, be used to create so-called *tensor densities* on arbitrary manifolds with a metric, and has fascinating applications in Hodge-de Rham theory in differential geometry.

CHAPTER 5

Method of stationary phase

This method was originally developed by Stokes and Kelvin in the 19th century in their study of water waves. Consider the general Laplace integral in which $\phi(t)$ is pure imaginary, i.e. $\phi(t) = i\psi(t)$ so that

$$I(x) = \int_a^b f(t) e^{ix\psi(t)} dt,$$

where $a, b, x, f(t)$ and $\psi(t)$ are all real. In this case, $I(x)$ is called the *generalized Fourier integral*. Note that the term $e^{ix\psi(t)}$ is purely oscillatory and so we cannot exploit the exponential decay of the integrand away from a maximum as was done in Watson's lemma and Laplace's method. However, if x is large, the integrand in $I(x)$ oscillates rapidly and so we might expect approximate cancellation of positive and negative contributions from adjacent intervals, leading to a small net contribution to the integral. In fact we have the following result.

RIEMANN-LEBESGUE LEMMA. *If $|f(t)|$ is integrable and $\psi(t)$ is continuously differentiable over $a \leq t \leq b$, but $\psi(t)$ is not constant over any subinterval of $a \leq t \leq b$, then*

$$I(x) = \int_a^b \overbrace{f(t) e^{ix\psi(t)}}^{u dv} dt \rightarrow 0 \quad \text{as } x \rightarrow +\infty.$$

→ To obtain an asymptotic expansion of $I(x)$ as $x \rightarrow +\infty$, we integrate by parts as before. This is valid provided the boundary terms are finite and the resulting integral exists; for example, provided that $f(t)/\psi'(t)$ is smooth over $a \leq t \leq b$ and non-zero at either boundary, then

$$I(x) = \left[\frac{f(t)}{ix\psi'(t)} \cdot \underbrace{e^{ix\psi(t)}}_v \right]_a^b - \frac{1}{ix} \int_a^b \frac{d}{dt} \left(\frac{f(t)}{\psi'(t)} \right) \underbrace{e^{ix\psi(t)}}_v dt. \quad \checkmark$$

The 2nd integral is $o(\frac{1}{x})$ as $x \rightarrow +\infty$ by the Riemann-Lebesgue lemma and so

$$I(x) \sim \left[\frac{f(t)}{ix\psi'(t)} \cdot e^{ix\psi(t)} \right]_a^b \quad \text{as } x \rightarrow +\infty.$$

But Integration by parts may fail if $\psi(t)$ has a *stationary point* in the range of integration, i.e. $\psi'(c) = 0$ for some $a \leq c \leq b$. Suppose that indeed $\psi'(c) = 0$ for some $a \leq c \leq b$ and $\psi'(t) \neq 0$ for all $t \neq c$ in $[a, b]$. Then proceed as follows:

$$\begin{aligned} \frac{du}{dt} &= \frac{1}{ix} \frac{d}{dt} \left(\frac{f}{\psi'} \right) \\ dv &= ix \psi' e^{ix\psi} \\ u dv &= f \cdot e^{ix\psi} \end{aligned}$$

ie, does not
diverge, a
crucial
assumption
often not
true!

Step 1. For a small $\epsilon > 0$ we decompose $I(x)$ to $I(x; \epsilon)$ just like in (??) for Laplace's method. The remainder terms we neglect, for example in the case when $a < c < b$, are

$$\int_a^{c-\epsilon} f(t) e^{ix\psi(t)} dt + \int_{c+\epsilon}^b f(t) e^{ix\psi(t)} dt$$

and these vanish like $1/x$ as $x \rightarrow +\infty$, as $\psi(t)$ has no stationary points in either interval, and we can integrate by parts and apply the Riemann-Lebesgue lemma.

Step 2. With ϵ small enough, in order to obtain the leading order behaviour, we replace $f(t)$ and $\psi(t)$ by the approximations

$$f(t) \approx f(c) \quad \text{and} \quad \psi(t) \approx \psi(c) + \frac{\psi''(c)}{2!} \cdot (t-c)^2.$$

This assumes $\psi''(c) \neq 0$, otherwise we must consider higher order terms. Hence we get

$$I(x) \sim \int_{c-\epsilon}^{c+\epsilon} f(c) e^{ix \cdot \left(\psi(c) + \frac{\psi''(c)}{2!} \cdot (t-c)^2 \right)} dt \quad \text{as } x \rightarrow +\infty.$$

Step 3. We replace ϵ by infinity—this again introduces terms which vanish like $1/x$ as $x \rightarrow +\infty$ and which will be asymptotically smaller terms that can be neglected. Then setting (a change of variables due to Morse)

$$s^2 = x \cdot \frac{|\psi''(c)|}{2!} \cdot (t-c)^2 \cdot \text{sgn}(\psi''(c))$$

yields

$$I(x) \sim f(c) e^{ix\psi(c)} \cdot \sqrt{\frac{2}{x|\psi''(c)|}} \cdot \int_{-\infty}^{\infty} e^{i \cdot \text{sgn}(\psi''(c)) \cdot s^2} ds \quad \text{as } x \rightarrow +\infty. \quad (5.1)$$

To evaluate the integral on the right, use that $\int_{-\infty}^{\infty} e^{\pm is^2} ds = \sqrt{\pi} e^{\pm i\pi/4}$ so that eventually we get

$$I(x) \sim f(c) e^{ix\psi(c) + i \cdot \text{sgn}(\psi''(c)) \pi/4} \cdot \sqrt{\frac{2\pi}{x|\psi''(c)|}} \quad \text{as } x \rightarrow +\infty.$$

Note. If $c = a$ or $c = b$, the contribution from the integral, which is only over a semi-infinite interval, means that the asymptotic result above must be multiplied by a factor of $\frac{1}{2}$. If $\psi(t)$ has many stationary points in $[a, b]$, then we split up the integral into intervals containing only one stationary point and deal with each one independently (though their contributions are additive). Again, the relative size of $f(t)$ at the stationary points of $\psi(t)$ will now be important. If the stationary point is such that $\psi'(c) = \psi''(c) = \dots = \psi^{m-1}(c) = 0$ and $\psi^m(c) \neq 0$ then make the approximation $\psi(t) \approx \psi(c) + \frac{\psi^m(c)}{m!} \cdot (t-c)^m$ instead. In this case $I(x)$ behaves like $x^{-1/m}$ as $x \rightarrow +\infty$.

Note. Alternatively, after (5.1), we could have made the substitution $s = v e^{i \cdot \text{sgn}(\psi''(c)) \pi/4}$ and used that $\int_{-\infty}^{\infty} e^{-v^2} dv = \sqrt{\pi}$.

Example. To find the leading order asymptotic behaviour of

$$I(x) = \int_0^1 \cos x(t^3 - t) dt \quad \text{as } x \rightarrow +\infty,$$

we first write the integral in the form

$$I(x) = \operatorname{Re} \left\{ \int_0^1 e^{ix(t^3 - t)} dt \right\}$$

so that we can identify $f(t) \equiv 1$ and $\psi(t) = t^3 - t$. Hence $\psi'(t) = 3t^2 - 1$ and $\psi''(t) = 6t$ so that ψ has two stationary points, of which only the positive one lies in the range of integration at $t_0 = 1/\sqrt{3}$. Since $\psi''(t) = 6t > 0$ for $t \geq 0$, this is a local and global minimum; and finally note that $\psi(t_0) = -2/(3^{3/2})$. Hence truncating our interval of integration to a small neighbourhood of t_0 (this only introduces asymptotically smaller terms), we then make the approximation

$$\psi(t) \approx \psi(t_0) + \frac{\psi''(t_0)}{2} \cdot (t - t_0)^2$$

and then extending the endpoints of integration off to infinity (again only introduces asymptotically smaller terms), we finally get

$$I(x) \sim e^{ix\psi(t_0)} \int_{-\infty}^{\infty} e^{ix \cdot \frac{\psi''(t_0)}{2} \cdot (t-t_0)^2} dt, \quad \text{as } x \rightarrow +\infty,$$

Making the substitution $s^2 = x \cdot \frac{\psi''(t_0)}{2} \cdot (t - t_0)^2$, generates

$$I(x) \sim e^{ix\psi(t_0)} \cdot \sqrt{\frac{2}{x\psi''(t_0)}} \cdot \int_{-\infty}^{\infty} e^{is^2} ds = e^{i(x\psi(t_0) + \pi/4)} \cdot \sqrt{\frac{2\pi}{x\psi''(t_0)}}$$

Now taking the real part of this, we get

$$I(x) \sim \sqrt{\frac{\pi}{x\sqrt{3}}} \cos \left(\frac{\pi}{4} - \frac{2x}{3^{3/2}} \right) \quad \text{as } x \rightarrow +\infty.$$

$$\int_0^\pi d\phi e^{\pm i a \cos(\phi - \phi')} = \int_0^\pi d\phi \cos(a \overset{a \sin}{\downarrow} \cos[\phi - \phi']) = \pi J_0(a)$$

↑
or sin

$$\int_0^a J_0(s) ds = a J_1(a) \quad J_0(0) = 1 \quad J_1(0) = 0$$

$$\int_0^a s J_0(s) ds = a J_1(a)$$

$$\lim_{x \rightarrow 0} \frac{J_1(x)}{x} = \frac{1}{2}$$

$$\int_0^a s J_0(bs) ds = a^2 \left[\frac{J_1(ab)}{(ab)} \right]$$

$$\int_0^a s J_0(bs) ds = s^2 \frac{J_1(bs)}{(bs)}$$

$$\int_0^1 s^{p+1}$$

$$\int_{-1}^1 e^{iax} = \frac{2 \sin a}{a}$$

$$\int_0^{\pi} \sin x dx = 2$$

SELECTED DEFINITE INTEGRALS

$$\int_0^{\pi} \sin^2(mx) dx = \int_0^{\pi} \cos^2(mx) dx = \frac{\pi}{2} \int_0^{\pi}$$

$$\int_0^{\infty} \frac{\sin^2 x}{x^2} dx = \frac{\pi}{2} \quad \int_0^{\infty} \frac{\sin mx}{x} dx = \begin{cases} \pm \frac{\pi}{2} & \begin{cases} + m > 0 \\ - m < 0 \end{cases} \\ 0 & m = 0 \end{cases}$$

$$\int_0^{\infty} e^{-\alpha x} dx = \frac{1}{\alpha} \quad (\alpha > 0)$$

$$\int_0^{\infty} e^{-\alpha x} x dx = \frac{1}{\alpha^2}$$

$$\int_0^{\infty} e^{-\alpha x} x^2 dx = \frac{2}{\alpha^3}$$

$$\int_0^{\infty} e^{-\alpha x} x^n dx = \frac{n!}{\alpha^{n+1}}$$

$$\Gamma(n) = (n-1)!$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} x^2 dx = \frac{1}{2\alpha} \sqrt{\frac{\pi}{\alpha}}$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} x^{2n} dx = \frac{(2n-1)!!}{(2\alpha)^n} \sqrt{\frac{\pi}{\alpha}}$$

$5!! = 5 \cdot 3 \cdot 1 \quad 6!! = 6 \cdot 4 \cdot 2$

$$\int \sin^2 \theta d\Omega = \frac{8\pi}{3}$$

$$\int \cos^2 \theta d\Omega = \frac{4\pi}{3}$$

$$\int_0^{\infty} e^{-\alpha x^2} x dx = \frac{1}{2\alpha}$$

$$\int_{-\infty}^{\infty} dz e^{ikz} = 2 \frac{\sinh k}{k}$$

$$\int_0^{\infty} e^{-\alpha x^2} x^{2n+1} dx = \frac{n!}{2 \alpha^{n+1}}$$

$$\int_{-\infty}^{\infty} e^{ik(z-z')} dk = 2\pi \delta(z-z')$$

$$\int_0^{\infty} e^{-\mu r} \sin qr dr = \frac{q}{\mu^2 + q^2}$$

$$\int_0^{\infty} e^{-\alpha x} \frac{dx}{\sqrt{x}} = \sqrt{\frac{\pi}{\alpha}}$$

$$\int_0^{\infty} e^{-x^2} x^{\alpha} dx = \frac{1}{2} \Gamma\left(\frac{\alpha+1}{2}\right)$$

More selected integrals

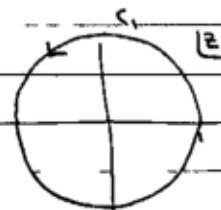
$$\textcircled{1} \int_0^{\infty} dx e^{-\mu x} x^{\nu-1} = \frac{\Gamma(\nu)}{[\mu]^{\nu}} \quad \Rightarrow \quad \int_1^{\infty} ds (s)^{-\alpha} (\ln s)^{\beta} = \frac{\Gamma(\beta+1)}{[\alpha-1]^{\beta+1}} \quad (\text{GR p317})$$

$$\Rightarrow \int_1^{\infty} ds s^{-j-1} [(s)^a (\ln s)^b] = \frac{\Gamma(b+1)}{(j-a)^{b+1}}$$

$$\textcircled{2} \int_0^{\infty} dx x^a (x-i)^b = B(-a-b-1, b+1) \quad / \text{p. 285 GR.}$$

$$\textcircled{3} \int d^3r e^{i\mathbf{k}\hat{\mathbf{a}}\cdot\vec{r}} \left(\frac{1}{r}\right)^{\frac{(V\hat{\mathbf{a}})}{2}} = \int d^3r e^{i\vec{k}\cdot\vec{r}} \left[\frac{1}{|\vec{r}|}\right] = \frac{4\pi}{|\vec{k}|^2}$$

$$\int_0^{2\pi} \frac{dx}{a+b\cos x} = ?$$



Let $e^{ix} = z$, $dz = iz dx$ $\cos x = \frac{1}{2}(z + \bar{z})$

$$\int_{\gamma} \frac{dz}{i z (a + \frac{1}{2}b(z + \bar{z}))} = \frac{1}{i} \int_{\gamma} \frac{dz}{az + \frac{1}{2}bz^2 + \frac{1}{2}b}$$

$$= \frac{1}{i} \frac{1}{(\frac{1}{2}b)} \int_{\gamma} \frac{dz}{[z^2 + 2\frac{a}{b}z + 1]} = \frac{1}{i(\frac{1}{2}b)} \oint \frac{dz}{(z-A)(z-B)}$$

where $A = -\frac{a}{b} - \sqrt{(\frac{a}{b})^2 - 1}$

$B = -\frac{a}{b} + \sqrt{(\frac{a}{b})^2 - 1}$

Where are these poles? Let $x = \frac{a}{b}$, so $-x \pm \sqrt{x^2 - 1}$.

Assume $x = \text{real}$. Then

$$|-x \pm \sqrt{x^2 - 1}|^2 = (-x \pm \sqrt{x^2 - 1})^2 = x^2 - 1 \mp 2x\sqrt{x^2 - 1}$$

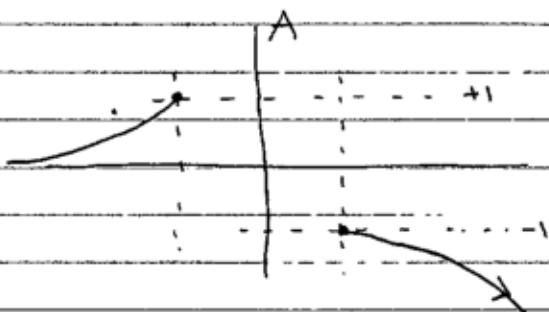
Well, let $x = z$, consider

$$A(z) = -z - \sqrt{z^2 - 1}$$

$$A'(z) = -1 - \frac{z}{\sqrt{z^2 - 1}}$$

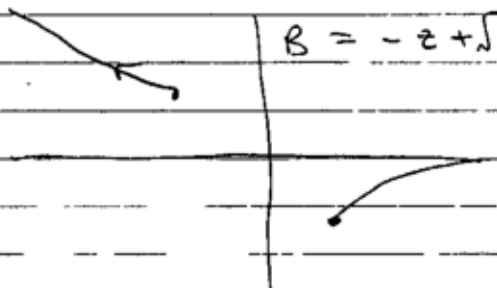
For $z > 1$, $A'(z) < 0$, so $\nearrow$.

Picture shows: for $z < -1$, this root lies inside circle. For $z > +1$, root lies outside



$$B = -z + \sqrt{z^2 - 1}$$

For $z > 1$, reverse is true! So, for $z > 1 \wedge z < -1$, exactly one root lies inside.



So, assume $z = \gamma = \frac{a}{b} > 1$. Then

$$\frac{1}{i \frac{1}{2}b} \oint_C \frac{dz}{(z-A)(z-B)} = \frac{1}{i(\frac{1}{2}b)} \oint_B \frac{dz}{(z-B)} \times \frac{1}{(B-A)}$$

$$= \frac{1}{(2\sqrt{\gamma^2-1}) \frac{1}{2}b} \cdot 2\pi \cdot \underbrace{\frac{1}{2\pi i} \oint \frac{dz}{z-B}}_{=1}$$

$$= \frac{2\pi}{b\sqrt{(\frac{a}{b})^2-1}} = \frac{2\pi}{\sqrt{a^2-b^2}} \quad \text{assume } \left(\frac{a}{b}\right) > 1$$

$\nwarrow$ a, b have same sign
and $|a| > |b|$.

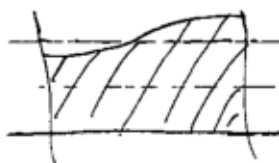
So: $\int_0^{2\pi} \frac{dx}{(a+b\cos x)}$

However, if $z = \gamma = \frac{a}{b} < -1$, then the root A is say let

$$\frac{1}{i(\frac{1}{2}b)} \oint_A \frac{dz}{z-A} \left(\frac{1}{A-B} \right) = \frac{+2\pi}{\sqrt{a^2-b^2}}$$

But $\frac{a}{b} < -1 \Rightarrow \frac{a}{b} = -|\frac{a}{b}|$ and $|\frac{a}{b}| > 1 \Rightarrow |a| > |b|$.

Logic: If $a > 0$ and $a > |b|$, surely result is true.



I would restrict real validity to $a > |b| > 0$.

Conclusions:

$$\int_0^{2\pi} \frac{dx}{a+b\cos x} = \frac{2\pi}{\sqrt{a^2-b^2}}$$

validity for real a, b : $a > |b| \geq 0$.

But periodic so go to $\int_{-\pi}^{\pi} = 2 \int_0^{\pi}$, so same rule.

So

$$\int_0^{\pi} dx$$

Conclusions so far:

$$\boxed{\int_0^{\pi} \frac{dx}{a+b\cos x} = \frac{1}{2} \int_0^{2\pi} \frac{dx}{a+b\cos x} = \frac{\pi}{\sqrt{a^2-b^2}}}$$

valid for $a > |b| \geq 0$ a, b real

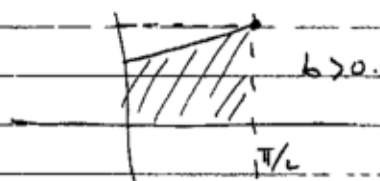
1. GR7 p405 $n=0$

2. Maple confirms $a > 0, b > 0, a > b$

$$\int_0^{\pi/2} \frac{dx}{a+b\cos x} = \frac{|\cos^{-1}(b/a)|}{\sqrt{a^2-b^2}}$$

where $0 < \cos^{-1} x \leq \pi$. Validity: $a > |b| > 0$. Easier to do as

$$\frac{2}{\sqrt{a^2-b^2}} \tan^{-1} \left[\sqrt{\frac{a-b}{a+b}} \right] \quad \text{where } -\frac{\pi}{2} < \tan^{-1} < \frac{\pi}{2}$$



$$\int_0^{\pi/2} \frac{dx}{a+b\cos x} = \frac{\theta}{\sqrt{a^2-b^2}}$$

where $\cos \theta = (b/a)$, $0 \leq \theta \leq \pi$

validity: $a > |b| > 0$, a, b real.

Next,

$$\int_0^{\pi} \frac{dx}{a+b\sin x}$$

$$= \int_{-\pi/2}^{\pi/2} \frac{dx}{a+b\sin[x+\pi/2]}$$

$$= \int_{-\pi/2}^{\pi/2} \frac{dx}{a+b\cos x} = 2 \int_0^{\pi/2} \frac{dx}{a+b\cos x} = \frac{2\theta}{\sqrt{a^2-b^2}}$$

So:

$$\int_0^{\pi} \frac{dx}{a+b\sin x} = \frac{2\theta}{\sqrt{a^2-b^2}}$$

where $\cos \theta = (b/a)$, $0 \leq \theta \leq \pi$

validity $a > |b| > 0$; a, b real.

$$\int_0^{\pi} \frac{dx}{a^2 - b^2 \cos^2 x} = \int_0^{\pi} dx \left[\frac{1}{a + b \cos x} + \frac{1}{a - b \cos x} \right] \left(\frac{1}{2a} \right)$$

$$= \frac{1}{2a} \cdot 2 \cdot \frac{\pi}{\sqrt{a^2 - b^2}} = \frac{\pi}{a \sqrt{a^2 - b^2}}, \quad a > |b| > 0.$$

$$\int_0^{\pi/2} \frac{dx}{a^2 - b^2 \cos^2 x} = \int_0^{\pi/2} dx \left[\frac{1}{a + b \cos x} + \frac{1}{a - b \cos x} \right] \left(\frac{1}{2a} \right)$$

$$= \frac{1}{2a} \left[\frac{\cos^{-1}(b/a)}{\sqrt{a^2 - b^2}} + \frac{\cos^{-1}(-b/a)}{\sqrt{a^2 - b^2}} \right]$$

$$\cos^{-1}\left(-\frac{b}{a}\right) = -\cos^{-1}\left(\frac{b}{a}\right) + \pi$$

i.e. $\cos^{-1}(x) + \cos^{-1}(-x) = \pi$. ✓



$$= \frac{1}{2a} \frac{\pi}{\sqrt{a^2 - b^2}} \left[\pi \right] = \frac{\pi}{2a \sqrt{a^2 - b^2}}, \quad a > |b| > 0.$$

$$\int_0^{\pi} \frac{dx}{a^2 - b^2 \sin^2 x} = \frac{1}{2a} \int_0^{\pi} dx \left[\frac{1}{a + b \sin x} + \frac{1}{a - b \sin x} \right]$$

$$= \frac{\pi}{2a \sqrt{a^2 - b^2}}, \quad \text{twice the above:}$$

$$\int_0^{\pi} \frac{dx}{a^2 - b^2 \cos^2 x} = \frac{\pi}{a \sqrt{a^2 - b^2}}, \quad a > |b| > 0$$

$$\int_0^{\pi} \frac{dx}{a^2 - b^2 \sin^2 x} = 2 \int_0^{\pi/2} \frac{dx}{a^2 - b^2 \cos^2 x} = \frac{\pi}{a \sqrt{a^2 - b^2}}$$

$\int$

$$\int_0^{\pi} \frac{dx}{a^2 - b^2 \cos^2 x} = \int_0^{\pi} \frac{dx}{a^2 - b^2 \sin^2 x} = \frac{\pi}{|a| \sqrt{a^2 - b^2}}$$

$$\int_0^{\frac{\pi}{2}} \frac{dx}{a^2 - b^2 \cos^2 x} = \frac{\pi}{2|a| \sqrt{a^2 - b^2}}$$

validity: $|a| > |b| \geq 0$ a, b real

About Integrals.

① Consider $\int_0^\pi d\theta f[\cos^2\theta, \sin^2\theta]$. Can rewrite as $\int_0^\pi d\theta \tilde{f}[\cos 2\theta]$,
so period is π . One can shift to $\int_{-\pi/2}^{\pi/2}$. Then use $\theta \rightarrow -\theta$ reflection
to get $2 \int_0^{\pi/2}$. So

$$\boxed{\int_0^\pi d\theta f[\cos^2\theta, \sin^2\theta] = 2 \int_0^{\pi/2} d\theta f[\cos^2\theta, \sin^2\theta]}$$

② Also, since $\cos^2[\theta + \pi/2] = \sin^2\theta$, get
 $\sin^2(\theta + \pi/2) = \cos^2\theta$

$$\boxed{\int_0^\pi d\theta f[\cos^2\theta, \sin^2\theta] = \int_0^\pi d\theta f[\sin^2\theta, \cos^2\theta]}$$

Gaussian Integrals

① Complete the square or look up Spiegel p98 to find:

$$\int_{-\infty}^{\infty} dx e^{-(ax^2+bx+c)} = \sqrt{\frac{\pi}{a}} \exp\left[-\frac{b^2-4ac}{4a}\right] \equiv I_0(a,b,c)$$

② Now:

$$\begin{aligned} I_n(a,b,c) &\equiv \int_{-\infty}^{\infty} dx x^n e^{-(ax^2+bx+c)} \\ &= (-1)^n (2/2b)^n I_0(a,b,c) = (-2/2b) I_{n-1}(a,b,c) \end{aligned}$$

In particular,

$$I_1(a,b,c) = - I_0(a,b,c) \times (b/2a)$$

$$I_2(a,b,c) = + I_0(a,b,c) \times \left\{ (1/2a) + (b/2a)^2 \right\}$$

Theorem: For a well-behaved finite $\vec{f}(\theta, \phi)$, periodic in ϕ ,
we have $\int d\Omega (\vec{\nabla} \cdot \vec{f}) = 0$

Proof: $\int d\Omega (\vec{\nabla} \cdot \vec{f}) = \int S_\theta d\phi \left[\frac{1}{r S_\theta} \partial_\theta (S_\theta f_\theta) + \frac{1}{r S_\theta} \partial_\phi (f_\phi) \right]$
where $\vec{f} = \hat{r} f_r + \hat{\theta} f_\theta + \hat{\phi} f_\phi$

Now consider integral as 2 integrals

$$\begin{aligned} (1) &= \frac{1}{r} \int d\phi \underbrace{\int d\theta \partial_\theta (S_\theta f_\theta)}_{\pi} \\ &= (S_\theta f_\theta)_{-\pi}^{\pi} = 0 \text{ since } f_\theta = \text{finite} \end{aligned}$$

$$\begin{aligned} (2) &= \frac{1}{r} \int d\theta \underbrace{\int d\phi \partial_\phi (f_\phi)}_{2\pi} \\ &= f_\phi \Big|_0^{2\pi} = 0 \text{ since periodic} \end{aligned}$$

Corollary:

$$\int \vec{\nabla} f \cdot \vec{\nabla} g d\Omega = - \int f (\nabla^2 g) d\Omega + \underbrace{\int \vec{\nabla} \cdot (f \vec{\nabla} g) d\Omega}$$

Think of $\vec{F} = f \vec{\nabla} g$, so $F_\theta = f \frac{1}{r} \partial_\theta g$.

Assuming $f, \partial_\theta g$ are finite at $\pm\pi$, this $\rightarrow 0$!

Therefore, for such functions

$$\int \vec{\nabla} f \cdot \vec{\nabla} g d\Omega = - \int f \nabla^2 g d\Omega$$

as if you were "doing parts" with the ∇ .

which can be verified by elementary integration. We now introduce the variables u_i

$$\begin{aligned} z_1 &= u_1, \\ z_2 &= u_1 u_2, \\ &\vdots \\ z_{n-1} &= u_1 u_2 \cdots u_{n-1}. \end{aligned} \quad (\text{A5-4})$$

The identity (A5-2) thereby attains another useful form,

$$\frac{1}{a_1 a_2 \cdots a_n} = (n-1)! \int_0^1 u_1^{n-2} du_1 \int_0^1 u_2^{n-3} du_2 \cdots \int_0^1 du_{n-1} \times \frac{1}{[a_1 u_1 \cdots u_{n-1} + a_2 u_1 \cdots u_{n-2}(1-u_{n-1}) + a_n(1-u_1)]^n} \quad (\text{A5-5})$$

The integrations over the new variables $u_1 \cdots u_{n-1}$ are interchangeable. Assume that the identity (A5-2) and therefore (A5-5) is true for n polynomials. Then we can write for $n+1$ polynomials,

$$\frac{1}{a_1 a_2 \cdots a_n a_{n+1}} = (n-1)! \int_0^1 u_1^{n-2} du_1 \cdots \int_0^1 du_{n-1} \frac{1}{A^n a_{n+1}}, \quad (\text{A5-6})$$

where A^n is the denominator in the integral of (A5-5). But this expression can be rewritten by use of the equation

$$\int_0^1 \frac{u^{n-1} du}{[Au + B(1-u)]^{n+1}} = \frac{1}{nA^n B}, \quad (\text{A5-7})$$

which can be proved by elementary methods. We find, with $u = u_0$, $B = a_{n+1}$, by differentiating (A5-7)

$$\frac{1}{a_1 a_2 \cdots a_n a_{n+1}} = n! \int_0^1 u_0^{n-1} du_0 \cdots \int_0^1 du_{n-1} \frac{1}{[Au_0 + a_{n+1}(1-u_0)]^{n+1}} \quad (\text{A5-8})$$

Except for the labeling of the auxiliary variables, this equation is identical with Eq. (A5-5) for $n+1$ polynomials. The identity (A5-2) is thereby established by induction.

A5-1 Convergent integrals. We shall assume in this section that the integrals to be evaluated are absolutely convergent. The following operations can then be carried out without restrictions. The case of divergent integrals is separately discussed in the following section.

APPENDIX A5

ON THE EVALUATION OF CERTAIN INTEGRALS

In Chapter 9 we encountered integrals over the intermediate states of the system which play an important role both in the discussion of divergences and in the computations of observable phenomena.

The most convenient method of evaluation for these integrals was given by Feynman.* This method involves the introduction of auxiliary scalar variables, so that the original integration over the four-momentum space of the virtual particles reduces to a standard type of integration which can easily be done. One is then left with the integration over the auxiliary variables, which forms the main task in the actual computations.

There are three advantages to this procedure. One is its explicit covariance, the second is its property to reduce the number of poles to two (see below), and the third is its applicability to all integrals which occur in the evaluation of S -matrix elements, including divergent integrals.

A typical integral is of the form

$$I = \int \frac{F(k) d^4 k}{a_1 a_2 \cdots a_n}, \quad (\text{A5-1})$$

where the a_i are polynomials of second degree in the four-vectors k_i , and $F(k)$ is a polynomial of degree n' .

The introduction of the auxiliary variables, which we shall call $z_1, z_2, \dots$, is carried out so that the denominator $a_1 \cdots a_n$ becomes the n th power of a second degree polynomial in k_i :

$$\frac{1}{a_1 a_2 \cdots a_n} = (n-1)! \int_0^1 dz_1 \int_0^1 dz_2 \cdots \times \int_0^1 dz_{n-1} \frac{1}{[a_1 z_{n-1} + a_2(z_{n-2} - z_{n-1}) + \cdots + a_n(1 - z_1)]^n} \quad (\text{A5-2})$$

A product of n polynomials requires therefore $n-1$ auxiliary variables. The identity (A5-2) can be proved by induction as follows. For $n=1$ no auxiliary variable needs to be introduced. For $n=2$, we have

$$\frac{1}{a_1 a_2} = \int_0^1 \frac{dz}{[a_1 z + a_2(1-z)]^2} \quad (\text{A5-3})$$

* R. P. Feynman, *Phys. Rev.* 76, 769 (1949).

Grozin book generalization:

$$\frac{1}{a_1^{n_1} a_2^{n_2} \cdots} = \frac{\Gamma(n_1 + n_2 + \cdots + n_k)}{\Gamma(n_1) \Gamma(n_2) \cdots \Gamma(n_k)} \int \delta(x_1 + x_2 + \cdots - 1) x_1^{n_1-1} x_2^{n_2-1} \cdots dx_1 dx_2 \cdots dx_k$$

$$[a_1 x_1 + a_2 x_2 + \cdots]^{n_1 + n_2 + \cdots + n_k}$$

→
see own

Keep!

$$\frac{1}{q_1^{n_1} \dots q_k^{n_k}} = \frac{1}{P(n_1) \dots P(n_k)}$$

$$\times \int d\alpha_1 d\alpha_2 \dots d\alpha_k e^{-\alpha_1 \alpha_1 - \alpha_2 \alpha_2 - \dots - \alpha_1^{n_1-1} \alpha_2^{n_2-1} \dots \alpha_k^{n_k-1}}$$

$$\text{choose } \eta = \alpha_1 + \dots + \alpha_k \quad 1 \leq l \leq k$$

$$\text{insert } \delta(\alpha_1 + \alpha_2 + \dots + \alpha_k - \eta) d\eta.$$

$$\text{Let } \alpha_i = \eta x_i$$

do the η integration

get the gen result!!!

Try with $k=2$.

Keep!

$$\frac{1}{q_1^{n_1} \dots q_k^{n_k}} = \frac{1}{\Gamma(n_1) \dots \Gamma(n_k)} \\ \times \int d\alpha_1 d\alpha_2 \dots d\alpha_k \exp \left(-\frac{a_1 \alpha_1 + a_2 \alpha_2 + \dots}{\alpha_1^{n_1-1} \alpha_2^{n_2-1} \dots \alpha_k^{n_k-1}} \right)$$

choose $\eta = \alpha_1 + \dots + \alpha_k \quad 1 \leq l \leq k$

insert $\delta(\alpha_1 + \alpha_2 + \dots + \alpha_k - \eta) d\eta$.

Let $\alpha_i = \eta x_i$

do the η integration

get the gen result!!! $\rightarrow$

Try with $k=2$.

[next](#) [previous](#)

Next: [The W.K.B. solutions](#) **Up:** [Electromagnetic wave propagation](#) in **Previous:** [Ray tracing in the](#)

Asymptotic series: A mathematical aside

It is often convenient to expand a function of the complex variable $f(z)$ in inverse powers of z :

$$f(z) = \phi(z) \left[A_0 + \frac{A_1}{z} + \frac{A_2}{z^2} + \cdots \right], \quad (890)$$

where $\phi(z)$ is a function whose behaviour for large values of z is known. If $f(z)/\phi(z)$ is singular as $|z| \rightarrow \infty$ then the above series diverges. Nevertheless, under certain circumstances, the series may still be useful.

The circumstance needed to make this possible is that the difference between $f(z)/\phi(z)$ and the first $n+1$ terms of the series be of order $1/z^{n+1}$, so that for sufficiently large z this difference becomes vanishingly small. More precisely, the series is said to represent $f(z)/\phi(z)$ *asymptotically*, that is

$$f(z) \simeq \phi(z) \sum_{p=0}^{\infty} \frac{A_p}{z^p}, \quad (891)$$

provided that

$$\lim_{|z| \rightarrow \infty} \left\{ z^n \left[\frac{f(z)}{\phi(z)} - \sum_{p=0}^n \frac{A_p}{z^p} \right] \right\} \rightarrow 0, \quad \forall n > 0 \quad (892)$$

In other words, for a given value of n , the first $n+1$ terms of the series may be made as close as desired to the ratio $f(z)/\phi(z)$ by making z sufficiently large. For each value of z and n there is an error of order $1/z^{n+1}$. Since the series actually diverges, there is an optimum number of terms in the series used to represent $f(z)/\phi(z)$ for a given value of z . Associated with this is an unavoidable error. As z increases, the optimal number of terms increases and the error decreases. $\square$

Consider a simple example. The exponential integral is defined

$$E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt. \quad (893)$$

The asymptotic series for this function can be generated via a series of partial integrations. For example,

$$E_1(x) = \frac{e^{-x}}{x} - \int_x^\infty \frac{e^{-t}}{t^2} dt. \quad (894)$$

Continuing this procedure yields

$$E_1(x) = \frac{e^{-x}}{x} \left[1 - \frac{1}{x} + \frac{2!}{x^2} - \frac{3!}{x^3} + \cdots + \frac{(-1)^n n!}{x^n} \right] + (-1)^{n+1} (n+1)! \int_x^\infty \frac{e^{-t}}{t^{n+2}} dt. \quad (895)$$

$f(x) = x e^x E_1(x)$

$S_n(x)$

$\left\{ + (-1)^{n+1} (n+1)! \int_x^\infty \frac{e^{-t}}{t^{n+2}} dt \right\} \leftarrow \text{residual}$

$S_0 \quad f(x) - S_n(x) = x e^x \{ \}$

The infinite series obtained by taking the limit $n \rightarrow \infty$ diverges, since the Cauchy convergence test yields

$$\lim_{n \rightarrow \infty} \left| \frac{u_{n+1}}{u_n} \right| = \lim_{n \rightarrow \infty} \left[\frac{n}{x} \right] \rightarrow \infty. \quad \checkmark \quad (896)$$

Note that two successive terms in the series become equal in magnitude for $n = x$, indicating that the optimum number of terms for a given x is roughly the integer nearest x . To prove that the series is asymptotic, we need to show that

$$\lim_{x \rightarrow \infty} x^{n+1} e^x (-1)^{n+1} (n+1)! \int_x^\infty \frac{e^{-t}}{t^{n+2}} dt = 0. \quad (897)$$

$\uparrow$
OOPS.

$= x^n [f(x) - S_n(x)] \rightarrow 0. \quad \text{LHS} <$

This immediately follows, since

$$\int_x^\infty \frac{e^{-t}}{t^{n+2}} dt < \frac{1}{x^{n+2}} \int_x^\infty e^{-t} dt = \frac{e^{-x}}{x^{n+2}}. \quad \text{So above} \sim \frac{1}{x} \quad (898)$$

$$\sim \text{even } f(x) - S_n(x) < (-1)^{n+1} (n+1)! \frac{1}{x^{n+1}}.$$

Thus, the error involved in using the first n terms is less than $(n+1)! e^{-x} / x^{n+1}$, which is exactly the next term in the series. We can see that as n increases, this estimate of the error first decreases and then increases without limit. In order to visualize this phenomenon more exactly, let

$f(x) = x \exp(x) E(x)$, and let

$$J_n(x) = \sum_{p=0}^n \frac{(-1)^p p!}{x^p} \quad (899)$$

be the asymptotic series representation of this function which contains $n + 1$ terms. Figure 16 shows the relative error in the asymptotic series $|f_n(x) - f(x)|/f(x)$ plotted as a function of the approximate number of terms in the series n for $x = 10$. It can be seen that as n increases the error initially falls, reaches a minimum value at about $n = 10$, and then increases rapidly. Clearly, the optimum number of terms in the asymptotic series used to represent $f(10)$ is about 10.

Figure 16: The relative error in a typical asymptotic series plotted as a function of the number of terms in the series

Too bad

Asymptotic series are fundamentally different to conventional power law expansions, such as

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (900)$$

This series representation of $\sin x$ converges absolutely for all finite values of x . Thus, at any x the error associated with the series can be made as small as is desired by including a sufficiently large number of terms. In other words, unlike an asymptotic series, there is no intrinsic, or unavoidable, error associated with a convergent series. It follows that a convergent power law series representation of a function is *unique* inside the domain of convergence of the series. On the other hand, an asymptotic series representation of a function is *not unique*. It is perfectly possible to have two different asymptotic series representations of the same function, as long as the difference between the two series is less than the intrinsic error associated with each series. Furthermore, it is often the case that *different* asymptotic series are used to represent the *same* single-valued analytic function in different regions of the complex plane.

For example, consider the asymptotic expansion of the confluent hypergeometric function $F(a, c, x)$. This function is the solution of the differential equation

$$x F'' + (c - x) F' - a F = 0 \quad (901)$$

which is analytic at $z = 0$ [in fact, $F(a, c, 0) = 1$]. Here, $'$ denotes d/dz . The asymptotic expansion of $F(a, c, z)$ takes the form:

$$\begin{aligned} \frac{\Gamma(a)\Gamma(c-a)}{\Gamma(c)} F(a, c, z) &\simeq \Gamma(c-a) z^{a-c} e^z [1 + O(1/z)] \\ &\quad + \Gamma(a) z^{-a} e^{-i\pi a} [1 + O(1/z)] \end{aligned} \quad (902)$$

for $-\pi < \arg(z) < 0$, and

$$\begin{aligned} \frac{\Gamma(a)\Gamma(c-a)}{\Gamma(c)} F(a, c, z) &\simeq \Gamma(c-a) z^{a-c} e^z [1 + O(1/z)] \\ &\quad + \Gamma(a) z^{-a} e^{i\pi a} [1 + O(1/z)] \end{aligned} \quad (903)$$

for $0 < \arg(z) < \pi$, and

$$\begin{aligned} \frac{\Gamma(a)\Gamma(c-a)}{\Gamma(c)} F(a, c, z) &\simeq \Gamma(c-a) z^{a-c} e^{-i2\pi(a-c)} e^z [1 + O(1/z)] \\ &\quad + \Gamma(a) z^{-a} e^{i\pi a} [1 + O(1/z)] \end{aligned} \quad (904)$$

for $\pi < \arg(z) < 2\pi$, etc. It can be seen that the expansion consists of a linear combination of two asymptotic series (only the first term in each series is shown). For $|z| \gg 1$, the first series is exponentially larger than the second whenever $\operatorname{Re}(z) > 0$. We say that the first series is *dominant* in this region, whereas the second series is *subdominant*. Likewise, the first series is exponentially smaller than the second whenever $\operatorname{Re}(z) < 0$. We say that the first series is subdominant and the second series is dominant in this region.

Consider a region in which one or other of the series is dominant. Strictly speaking, it is not mathematically consistent to include the subdominant series in the asymptotic expansion because its contribution is actually less than the intrinsic error associated with the dominant series [this error is $O(1/z)$ times the dominant series, since we are only including the first term in this series]. Thus, at a general point in the complex plane the asymptotic expansion simply consists of the dominant series. However, this is not the case in the immediate vicinity of the lines $\operatorname{Re}(z) = 0$: these are called the *anti-Stokes lines*. When an anti-Stokes line is crossed, a dominant series becomes subdominant and *vice versa*. In the immediate vicinity of an anti-Stokes line neither series is dominant, so it is mathematically consistent to include both series in the asymptotic expansion.

The hypergeometric function $F(a, c, z)$ is a perfectly well behaved, single-valued, analytic function in the complex plane. However, our two asymptotic series are, in general, multi-valued functions in the complex plane [see Eq. (4.252a)]. Can a single-valued function be represented asymptotically by a multi-valued function? The short answer is no. We have to employ different combinations of the two series in different regions of the complex plane in order to ensure that $F(a, c, z)$ remains single-valued. Equations (4.252) show how this is achieved. Basically, the coefficient in front of the subdominant series changes *discontinuously* at certain values of $\arg(z)$. This is perfectly consistent with $F(a, c, z)$ being an analytic function because the subdominant series is "invisible"; i.e., the contribution of the subdominant series to the asymptotic solution falls below the intrinsic error associated with the dominant series, so it does not really matter if the coefficient in front of the former series changes discontinuously. Imagine tracing a large circle, centred on the origin, in the complex plane. Close to an anti-Stokes line, neither series is dominant, so we must include both series in the asymptotic expansion. As we move away from the anti-Stokes line, one series becomes dominant, which means that the other series becomes subdominant and, therefore, drops out of our asymptotic expansion. Eventually, we approach a second anti-Stokes line, and the subdominant series reappears in our asymptotic expansion. However, the coefficient in front of the subdominant series when it reappears is different to that which it had when it disappeared. This new coefficient is carried across the second anti-Stokes line into the region where the subdominant series becomes dominant. In this new region, the dominant series becomes subdominant and disappears from our asymptotic expansion. Eventually, a third anti-Stokes line is approached and the series reappears, but, again, with a different coefficient in front. The jumps in the coefficients of the subdominant series are chosen in such a manner that if we perform a complete circuit in the complex plane then the value of the asymptotic expansion is the same at the beginning and the end points. In other words, the asymptotic expansion is single-valued, despite the fact that it is built up out of two asymptotic series which are not single-valued. The jumps in the coefficient of the subdominant series, which are needed to keep the asymptotic expansion single-valued, are called *Stokes phenomena*, after the celebrated nineteenth century British mathematician Sir George Gabriel Stokes, who first drew attention to this effect.

Where exactly does the jump in the coefficient of the subdominant series occur? All we can really say is "somewhere in the region between two anti-Stokes lines where the series in question is subdominant." The problem is that we only retain the first term in each asymptotic series. Consequently, the intrinsic error in the dominant series is relatively large and we lose track of the subdominant series very quickly after moving away from an anti-Stokes line. However, we could include more terms in each asymptotic series. This would enable us to reduce the intrinsic error in the dominant series and, thereby, expand the region of the complex plane in the vicinity of the anti-Stokes lines where we can see both the dominant and subdominant series. If we were to keep adding terms to our asymptotic series, so as to minimize the error in the dominant solution, we would eventually be forced to conclude that a jump in the coefficient of the subdominant series can only take place on those lines in the complex plane on which $\text{Im}(z) = 0$: these are called *Stokes lines*. This result was

first proved by Stokes in 1857.¹⁵ On a Stokes line the magnitude of the dominant series achieves its *maximum* value with respect to that of the subdominant series. Once we know that a jump in the coefficient of the subdominant series can only take place at a Stokes line, we can retain the subdominant series in our asymptotic expansion in all regions of the complex plane. What we are basically saying is that, although, in practice, we cannot actually see the subdominant series very far away from an anti-Stokes line because we are only retaining the first term in each asymptotic series,

we could, in principle, see the subdominant series at all values of $\arg(z)$ provided that we retained a sufficient number of terms in our asymptotic series.

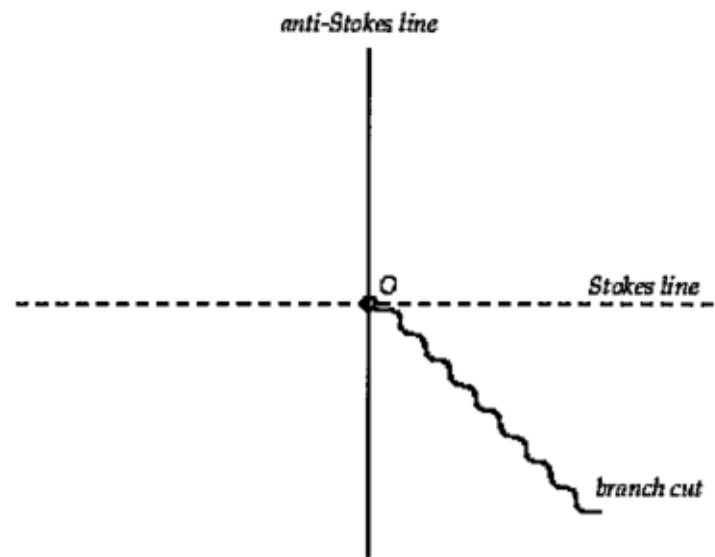


Figure 17: The location of the Stokes lines (dashed), the anti-Stokes lines (solid), and the branch cut (wavy) in the complex plane for the asymptotic expansion of the hypergeometric function

Figure 17 shows the location in the complex plane of the Stokes and anti-Stokes lines for the asymptotic expansion of the hypergeometric function. Also shown is a branch cut, which is needed to make z single-valued. The branch cut is chosen such that $\arg(z) = 0$ on the positive real axis. Every time we cross an anti-Stokes line the dominant series becomes subdominant and *vice versa*. Every time we cross a Stokes line the coefficient in front of the dominant series stays the same, but that in front of the subdominant series jumps discontinuously [see Eqs. (4.252)]. Finally, the jumps in the coefficient of the subdominant series are such as to ensure that the asymptotic expansion is single-valued.

[n](#)
[prev](#)

Next: [The W.K.B. solutions as](#) **Up:** [Electromagnetic wave propagation in](#) **Previous:** [Ray tracing in the](#)

Richard Fitzpatrick 2002-05-18

11 AUG 78

checks + repair 1-14-05,
27 years later!Gaussian Quadrature.

Theorem: suppose you have a polynomial $f(x)$ of degree $2n-1$. Then the integral of this function over an interval (a,b) against a weight $w(x)$ is given exactly by this formula:

$$\int_a^b dx w(x) f(x) = \sum_{i=1}^n A_i f(x_i) \quad \checkmark$$

Ie, you get the integral exactly by adding up n numbers. The function must be evaluated at n magic points (not $2n-1$ points). The x_i are the zeros of the polynomial $\phi_n(x)$ which is determined by the weight function and interval (a,b) . The A_i , called weights, are given by the formula:

$$A_i = h_{n-1} \cdot \left(\frac{k_n}{k_{n-1}} \right) \cdot \frac{g_2(x_i)}{g_0(x_i)} \cdot \frac{1}{[\phi_{n-1}(x_i)]^2} \quad \checkmark \quad \left(\frac{g_2}{g_0} = \frac{X(x_i)}{\beta_n} \right)$$

← ASP783

This formula is discussed and derived below.

Proof of the Theorem: Trivial. All you do is write the polynomial $f(x)$ in Lagrangian form and at once you get that $A_i = \text{INT}(w(x)L_i(x), x)$. By choosing the x_i in $L_i(x)$ to be the zeros of $\phi_n(x)$, you make the evaluation of A_i very easy, yielding the above simple formula.

This is, however, only a proof for $f(x)$ being a polynomial of degree n . Since the entire polynomial is specified by n numbers, it is not very surprising that the integral can be found by evaluating the function at n places. Far more interesting is the fact that the formula is exact even if $f(x)$ is raised to be a polynomial of degree $2n-1$. What makes that true? My way of seeing this is not the best but here it is anyway: Certainly everything is OK if $f(x)$ is of degree n . Imagine then that you have found the weights A_i and the zeros x_i using the above derivation. Then if you expand $f(x)$ into its "powers", here is what you have in fact shown:

$$f(x) = \sum_{j=0}^n b_j \phi_j(x) \quad \checkmark \quad \rightarrow \quad \delta_{j,0} = \sum_{i=1}^n A_i \phi_j(x_i) \quad (*)$$

ie, we know now that $(*)$ is valid for $j=0,1,2,\dots,n$. If you like, this can be taken as the equation you solve for the A_i in a matrix sense. If we can show that $(*)$ is in fact valid for j in the extended range $(0, 2n-1)$, then we have proven what we want.

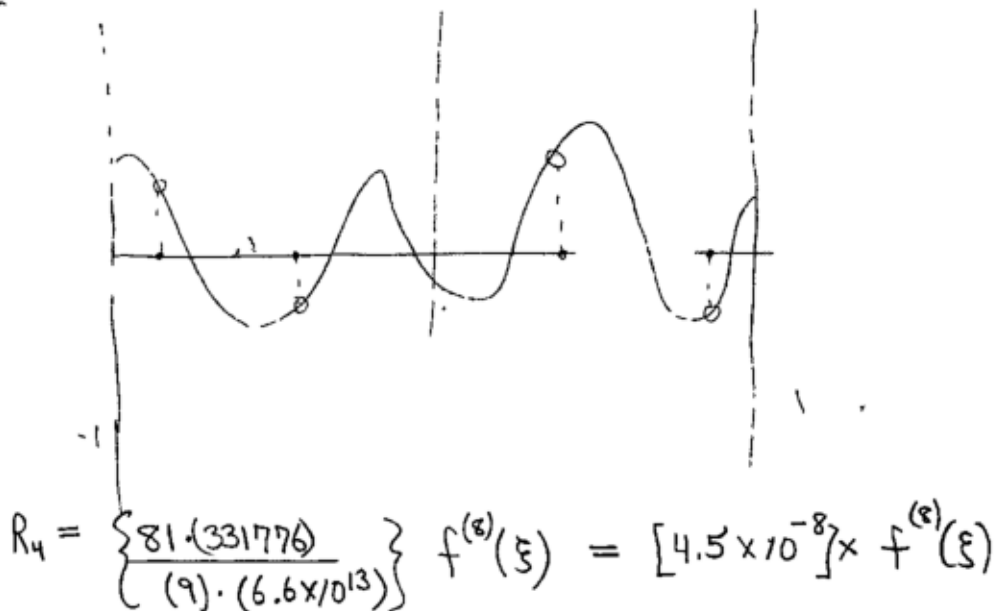
$$\sum_i A_i = u_0$$

because then we just multiply both sides by b_j and do the extended sum to conclude that the integral formula is exact for a polynomial of degree $2n-1$.

The proof that formula (*) can be extended in this way is given below. It is based on the recursion relation for the orthogonal functions.

Comments: thus we have a very powerful result. By knowing the value, say, of a 7th degree polynomial at only 4 points, we know the integral exactly. There are of course infinitely many 7th degree polynomials that run through these four points, but they must all have the same integral. Notice that these are not 4 arbitrary points, but are in fact 4 very special points.

Accuracy: Suppose we work with the $n=4$ Gauss-Legendre formula. How do we estimate the accuracy of an integral? The accuracy is perfect for polynomial of degree up to and including 7. Where are the gaussian points?



This last shows an estimate of the remainder. You see that if the function has no 8th derivative formula is exact. BUT if function has derivatives that tend to increase in size as you go to higher derivatives, then your error may be important. The thing about a polynomial is that its derivatives keep getting smoother. Suppose you apply to an 8th order polynomial. The 8th derivative is perhaps of order 1, more likely of order $8! = 40,000$, so error is at most 2 parts in 1000.

An obvious drawback of the gaussian formulas is that you cannot easily increase n as you can in Simpson's rule to see if result is stable.

Subroutine Extending the sum rule from m to $2m-1$.

① Suppose we know that, for $n = 0, 1, 2, \dots, m$, this is true:

$$C \cdot \delta_{n,0} = \sum_{i=1}^m w_i \phi_n(x_i)$$

where x_i are the m zeroes of ϕ_m . Claim: this thing is also true for $n = m+1, m+2, \dots, 2m-1$. It is not true for $n = 2m$ or higher. How to prove this?

② Symbolics:

$$0 = c$$

$$1 = 0$$

$$2 = 0$$

$$\vdots$$

$$m = 0$$

$$1 \otimes m = \sum_{i=1}^m w_i \phi_m(x_i) = \langle x, \phi_m(x) \rangle = (x, \phi_m(x)) \approx (p_0 + p_1, p_m) \sim \phi$$

$$\left(\sum_{i=1}^m w_i \phi_2(x_i) = 0. \right) \quad (2)$$

Show $(m+1) = 0$. Recursion says $(m+1) = x(m) + (m) + (m-1)$, can write this as $(m+1) = 1 \otimes m + (m) + (m-1)$. I.e., factor x acts like spin 1 since this is true $\uparrow$, i.e., vector addition.

But, whenever you see (m) , kill it because $\phi_m(x_i) \equiv 0$. Thus, $(m+1) = (m-1)$ and $(m-1) = 0$ since $\angle m \Rightarrow (m+1) = 0$.

Next, consider $(m+2)$. $(m+2) = 1 \otimes (m+1) + (m) + (m)$. Kill (m) and $(m+1)$, so $= 1 \otimes (m+1)$. But this is $1 \otimes [1 \otimes m + m + (m-1)] \Rightarrow 2 \otimes m + 1 \otimes (m-1)$. I keep only the significant terms. But $m=0$ and $1 \otimes (m-1) = m, m-1, m-2 \equiv 0$, so ok.

Go all the way to $2m-1$:

$$\begin{aligned} (2m-1) &= 1 \otimes (2m-2) + (2m-3) \\ &\equiv 2 \otimes (2m-3) + 1 \otimes (2m-4) \\ &\equiv 3 \otimes (2m-4) + 2 \otimes (2m-5) \\ &\vdots \\ &\equiv (m-1) \otimes m + (m-2) \otimes (m-1) \end{aligned}$$

comments: note first col Σ to $(2m-1)$, second col Σ to $(2m-3)$. First col reaches m before second col reaches $N \otimes N$, so OK. $N \otimes N$ is double because contains zero and $0 = c$, not 0 !

Now look at $(2m)$ + see that it fails

$$\begin{aligned} 2m &= 1 \otimes (2m-1) + (2m-2) \\ &\equiv 2 \otimes (2m-2) + 1 \otimes (2m-3) \\ &\equiv 3 \otimes (2m-3) + 2 \otimes (2m-4) \\ &\vdots \\ &\equiv m \otimes m + (m-1) \otimes (m-1) \end{aligned}$$

here is cause of the problem. For $n > 2m$, you will always reach $N \otimes N$ before you get to m .

③ An case the above is incomprehensible, here is an example.

Let $m=2$. Then recursion says:

$$\phi_3 = x \cdot \overset{=0}{\phi_2} + \overset{=0}{\phi_2} + \overset{=0}{\phi_1} \quad \Rightarrow \quad x\phi_2 = 0 \quad \checkmark$$

$$\Rightarrow \cancel{x(\cancel{x\phi_1} + \cancel{\phi_1} + \phi_0)}$$

$$\phi_4 = x \cdot \overset{=0}{\phi_3} + \overset{=0}{\phi_3} + \overset{=0}{\phi_2} \quad \Rightarrow \quad x\phi_3$$

$$\begin{aligned} &= x \left[\underset{=0}{x\phi_2} + \underset{=0}{\phi_2} + \phi_1 \right] = \underbrace{x\phi_1}_{\substack{\text{was called} \\ 1 \otimes 1}} = \phi_0 + \phi_1 + \phi_2 \\ &\quad \quad \quad \uparrow \neq 0. \end{aligned}$$

$$= \underline{2 \otimes 2} + 1 \otimes 1$$

Thus,

$$\sum_{i=1}^{m=2} w_i \phi_4(x_i) \neq 0!$$

Formula for the weights in general Gaussian integration.

1, You can get all these weights by just inverting a certain matrix whose elements are various functions f_n evaluated at the zeros of one of the functions. It is better and more direct though to use the formula derived here. }

2. The notation is a combination of that of Scheid and Abram-Stegun. I will try to do everything generally for all the orthogonal polynomial systems covered in Chapter 22 of the AS book.

In particular, here are some symbols I will use:

n = the order you are choosing. This is the number of zeros $x_1, \dots, x_n$, and the number of weights A_i . The x_i are zeros of the polynomial $\phi_n(x)$. $w(x)$ is the weight function, and the interval is (a, b) .

c_n = defined on page 773

k_n = defined on page 773, given p. 775

h_n = defined on page 774

= coeff of x^n in $\phi(n)$

g_2 and g_0 refer to table on page 783 which gives $d\phi_n/dx$. ✓

$PI(x)$ means $(x-x_1)(x-x_2)\dots(x-x_n)$ where x_i are the n zeros of $\phi(x)$. ✓

$L_1(x)$ is defined on page 53 of Num Anal Schwam. This is Lagrangian multiplier function which allows you to write a polynomial $p_n(x)$ of degree n in terms of the value of the polynomial at n arbitrary points $p_n(x_i)$. Note that there is no x_0 in our present usage, but there is on page 53. ✓

3: Derivation:

(a) First, you have to realize that the A_i are given by trivial integrals of the Lagrange function. This is completely obvious:

$$\int_a^b dx w(x) f(x) = \sum_{i=1}^n A_i f(x_i) \quad \text{/p. 125} \quad \checkmark$$

$$\text{But } f(x) = \sum_{i=1}^n L_i(x) f(x_i) \quad \text{/p. 53} \quad \checkmark$$

Compare to conclude that

$$A_i = \int_a^b dx w(x) L_i(x) \quad \checkmark$$

(b). Next, you observe that the functions Π and Π' which can be used to give $L_2(x)$ are simply related to ϕ_n as follows:

$$\phi_n(x) = k_n \cdot \Pi(x) \quad \checkmark$$

$$\Rightarrow \phi'_n(x) = k_n \Pi'(x) \quad \checkmark = \quad \phi'_n(x_i) = k_n \Pi'(x_i) \quad \checkmark$$

But from p 283, ^(22.8) $\phi'_n(x_i) = \underbrace{[g_0(x_i)/g_2(x_i)]}_{=(\beta_n/x)} \phi_{n+1}(x_i) \quad \checkmark \quad \phi_n(x_i) = 0 \quad \checkmark$

Thus:

$$\Pi'(x_i) = [g_0(x_i)/g_2(x_i)] \phi_{n+1}(x_i) / k_n \quad \checkmark$$

$$\Pi(x) = \phi_n(x) / k_n \quad \checkmark$$

(c) Using the form of $L_2(x)$, we then write A_i as integral of $\phi_n(x)/(x-x_i)$ like so:

$$L_i(x) = \frac{\Pi(x)}{\Pi'(x_i)(x-x_i)} \quad \checkmark = \frac{\phi_n(x)}{(x-x_i)} \cdot \frac{(g'_2/g'_0)}{\phi_{n+1}(x_i)} \quad \checkmark$$

$$\Rightarrow A_i = \frac{(g'_2/g'_0)}{\phi_{n+1}(x_i)} \int_a^b \frac{dx w(x) \phi_n(x)}{(x-x_i)} \quad \checkmark$$

(d). The next problem is to figure out this integral. The trick is to take the Christoffel-Darbous formulas on page 785 and specialize to $y=x_i$. You get:

$$\sum_{m=0}^n (\phi_m(x) \phi_m(x_i) / h_m) = \left(\frac{-k_n}{k_{n+1} h_n} \right) \phi_{n+1}(x_i) \cdot \left(\frac{\phi_n(x)}{x-x_i} \right) \quad \checkmark$$

We can simplify the RHS by replacing ϕ_{n+1} by ϕ_{n-1} according to 22.1.4, since x_i is a zero of ϕ_n . Thus:

$$\left(\frac{-k_n}{k_{n+1} h_n} \right) \phi_{n+1}(x_i) = \left(\frac{k_{n-1}}{k_n} \right) \left(\frac{\phi_{n-1}(x_i)}{h_{n-1}} \right) \quad \checkmark \quad \phi_{n+1} = -c_n \phi_{n-1}$$

$$\Rightarrow \text{RHS} = \frac{(k_{n-1}/k_n) \cdot \phi_{n-1}(x_i)}{h_{n-1}} \cdot \frac{\phi_n(x)}{(x-x_i)} \quad \checkmark$$

If you now integrate both sides of this sum rule against the weight $w(x)$, only the $m=0$ term survives because of orthogonality. ✓ Orthog generates a h_0 which cancels the h_0 on the left so result is:

$$1 = \frac{\phi_0(x_i)}{\phi_0} = \left(\frac{k_{n-1}}{k_n} \right) \frac{\phi_{n-1}(x_i)}{h_{n-1}} \int_a^b dx w(x) \frac{\phi_n(x)}{(x-x_i)} \quad \checkmark$$

↖ repair!

But ϕ_0 is a poly of degree zero, ie, a constant, so call it $\phi_0(1)$. These you can read off the table on page 774, called standardization. Thus, when x_i is one of the zeros of ϕ_n you may conclude that:

$$\int_a^b dx w(x) \frac{\phi_n(x)}{(x-x_i)} = \frac{h_{n-1} \cancel{\phi_{n-1}(x_i)}}{\phi_{n-1}(x_i)} \left(\frac{k_n}{k_{n-1}} \right) \quad \checkmark$$

Thus:

$$A_i = \cancel{\phi_{n-1}(x_i)} \cdot h_{n-1} \cdot \left(\frac{k_n}{k_{n-1}} \right) \cdot \left(\frac{g_2(x_i)}{g_0(x_i)} \right) \cdot \frac{1}{[\phi_{n-1}(x_i)]^2} \quad \checkmark \quad 1/05$$

$\equiv F$, see table $\rightarrow$

$$= \frac{\phi_{n-1}(x_i) A_i}{F \cdot (g_2/g_0)}$$

Table

see p777 top.

$\phi_n(x)$	$w(x)$	(a,b)	$\phi_0(1)$	k_n	h_{n-1}	g_2	g_0	k_n/k_{n-1}
$P_n(x)$	1	$(-1,1)$	1	$\frac{(2n)!}{2^n(n!)^2}$	$\frac{2}{2n-1}$	$(1-x_i^2)$	n	$(2n-1)/n$
$T_n(x)$	$\frac{1}{\sqrt{1-x^2}}$	$(-1,1)$	1	2^{n-1}	$\begin{cases} \frac{\pi}{2}, n \geq 1 \\ \pi, n=0 \end{cases}$	$(1-x_i^2)$	n	2
$U_n(x)$	$\sqrt{1-x^2}$	$(-1,1)$	$(1-x_i^2)$	2^n	$\frac{\pi}{2}$	$(1-x_i^2)$	$n+1$	2
$C_n^{(\alpha)}(x)$	$(1-x^2)^{\alpha-\frac{1}{2}}$	$(-1,1)$	$(1-x_i^2)$	$\frac{2^n \Gamma(\alpha+n)}{n! \Gamma(\alpha)}$	$\frac{\pi 2^{\alpha-1} \Gamma(\alpha)}{\Gamma(\alpha)^2}$	$(1-x_i^2)$	$n+2\alpha-1$	
$H_n(x)$	e^{-x^2}	$(-\infty, \infty)$	1	2^n	$\frac{\pi^{n/2}}{2^{n/2} (n-1)!}$	1	$2n$	
$L_n^{(\alpha)}(x)$	$e^{-x} x^\alpha$	$(0, \infty)$	1	$(-1)^n/n!$	$\frac{\Gamma(\alpha+n)}{(n-1)!}$	x_i	$(-n-\alpha)$	

ϕ_n	F
$P_n(x)$	$2(1-x_i^2)/n^2$
$T_n(x)$	$\pi(1-x_i^2)/n \quad (n \geq 2)$
$U_n(x)$	$\pi(1-x_i^2)/(n+1)$
$C_n^{(\alpha)}(x)$	
$H_n(x)$	
$L_n^{(\alpha)}(x)$	

$$p_i = k_i x + k_i'$$

$$x = \frac{p_i}{k_i} - \frac{k_i'}{k_i}$$

4-9

$$(x, p_m) = \left(\frac{p_i}{k_i}, p_m \right) - \frac{k_i'}{k_i} (1, p_m) = 0!$$

First-kind Gauss-Chebyshev Quadrature formulas:

① From our table we learned that

$$A_i = \frac{\pi}{n} \frac{(1-x_i^2)}{T_n'(x_i)^2}$$

But, $T_n(x) = \cos[n \cos^{-1}x]$ so zeros occur when $n \cos^{-1}x = \frac{\pi}{2}(2i-1)$

i.e., $x_i = \cos \alpha_i$ where $\alpha_i = \pi(2i-1)/2n$ $i=1, 2, \dots, n$. Then

$$T_n'(x_i) = \cos[(n-1)\alpha_i] = \cos[n\alpha_i - \alpha_i]. \text{ But } n\alpha_i = \frac{\pi}{2}(2i-1),$$

$$\text{so } \cos^2[n\alpha_i - \alpha_i] = \sin^2 \alpha_i. \text{ Hence } \boxed{A_i = \pi/n}$$

So

$$\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} f(x) = \frac{\pi}{n} \sum_{i=1}^n f(x_i)$$

$$\text{where } x_i = \cos \alpha_i \quad \alpha_i = \left(\frac{2i-1}{2n}\right)\pi \quad i=1, n.$$

Gauss-Legendre Quadrature Formulas:

① The zeros x_i are of course symmetric about $x=0$ since $P_n(x) = (-1)^n P_n(-x)$. But the zeros must be read from Tables on pp 916 - 919. Once you know the zeros, you may compute

②
$$A_i = \frac{2(1-x_i^2)}{n^2 P_n'(x_i)^2} = \text{the weights.}$$

③ Then:

$$\int_{-1}^1 dx f(x) = \sum_{i=1}^n A_i f(x_i)$$

Comments: I think this is the most famous Gaussian integration scheme. A slight problem is that you have to look up or compute the weights and zeros, since there do not exist closed formulas for these things as occurs in the Chubby cases. But these are listed and you only have to enter them once in a program to do such integration.

Accuracy: see elsewhere.

A Four Point Gauss-Legendre Program.

① You want to integrate

$\int_a^b dy F(y)$. Let $(a,b) \rightarrow (-1,1)$ by a

linear extension, so

$$x = \frac{2}{(b-a)} y + \frac{(a+b)}{(a-b)} = \frac{(2y-a-b)}{(b-a)} \quad \checkmark$$

$$\Rightarrow dx = \frac{2}{(b-a)} dy \quad \checkmark$$

$$\Rightarrow I = \frac{(b-a)}{2} \int_{-1}^1 dx f(x) \quad \checkmark$$

$$f(x) \equiv F\left(\underbrace{\left(\frac{b-a}{2}\right)x + \left(\frac{b+a}{2}\right)}_{y}\right) \quad \checkmark$$

$$\text{where } y = \frac{(b-a)}{2} x + \frac{(b+a)}{2} \quad \checkmark$$

$$\text{Let } c = (b-a)/2 \quad \checkmark, \text{ then } d = (b+a)/2. \quad I = c \cdot \sum_{i=1}^4 A_i F(cx_i + d) \quad \checkmark$$

$$\Rightarrow I = c \cdot A_1 (F(cx_1 + d) + F(-cx_1 + d)) + c \cdot A_2 (F(cx_2 + d) + F(-cx_2 + d)) \quad \checkmark$$

$\begin{matrix} A_2 c - x_1 \\ A_1 c + x_1 \end{matrix} \quad A_1 = A_3$
 $\uparrow$ sym.

So here is the program:

data A_1, A_2, x_1, x_2

get a, b and $f(x)$

$$c = (b-a)/2$$

$$d = (b+a)/2$$

$$I = c \cdot A_1 [f(d+cx_1) + f(d-cx_1)]$$

$$+ c \cdot A_2 [f(d+cx_2) + f(d-cx_2)]$$

page 916 A+S

$$A_1 = .65214 \quad 51548 \quad 62546$$

$$x_1 = .33998 \quad 10435 \quad 84856$$

$$A_2 = .34785 \quad 48451 \quad 37454$$

$$x_2 = .86113 \quad 63115 \quad 94053$$

Integration by Gaussian Quadratures

source: Miklos G.

application: Suppose you want to integrate $\int_{-1}^1 f(x) dx$ where $f(x)$ is some (known) function which is thought to be "as smooth as" a polynomial of order $2m-1$. Then here is how you can do the integral numerically:

$$\int_{-1}^1 f(x) dx = \sum_{i=1}^m w_i f(x_i) = I_m \quad (1)$$

where the $\{x_i\}$ are the m zeroes of $P_m(x)$, and where the $\{w_i\}$ are the constants which solve:

$$\delta_{no} = \sum_{i=1}^m w_i P_n(x_i) \quad n \in (0, m) \quad (2)$$

Here I_m should be a good estimate; you had to look up $2m$ numbers (the w_i and x_i). To check your answer, calculate I_{m+1} (you need $2(m+1)$ new numbers now). Claim is that convergence is extremely rapid.

Theory: it is clear how you solve (2) for the w_i . Using the Legendre recursion relations you can show that the relation is true in the larger range $n \in (0, 2m-1)$. Thus, you can imagine some zero weights + extend:

$$\delta_{no} = \sum_{i=1}^{2m-1} w_i P_n(x_i) \quad n \in (0, 2m-1) \quad w_i = 0 \text{ for } i \in (m+1, 2m-1)$$

Since the w_i and x_i satisfy this equation we can do $\sum_n b_n$ where b_n are coeffs of P_n in f , where f is approx by poly of order $2m-1$. Hence. and we know that $2m-1$

$$\int_{-1}^1 \underbrace{\left[\sum_n b_n P_n(x) \right]}_{f(x)} dx = \sum_{i=1}^{2m-1} w_i \underbrace{\left[\sum_n b_n P_n(x_i) \right]}_{f(x_i)} \quad \begin{array}{l} \text{points } f(x_i) \\ \text{completely sets} \\ \text{the integral.} \end{array}$$

Summation of Series (Miklos)

Suppose you want to do this:

$$\zeta(1.5) = \sum_{n=0}^{\infty} \frac{1}{n^{1.5}}$$

To get 13 decimals 18. accuracy, error must be $< 10^{-13}$; add of N terms so

$$\text{error} \approx \int_N^{\infty} \frac{1}{n^{1.5}} dn \sim \frac{1}{\sqrt{N}}$$

Thus, need

$$\text{err} = \frac{1}{\sqrt{N}} < 10^{-13} \Rightarrow N > 10^{26}$$

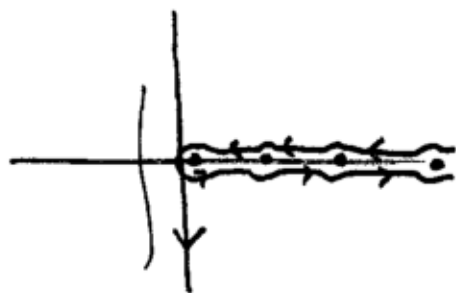
You would have to add 10^{26} terms. Oint no way.

The gimmick is this: we want to compute, say,

$$A = \sum_{n=0}^{\infty} f_n$$

Replace with contour:

$$A \approx \oint \frac{f(n) dn}{\tan(\pi n)}$$



up to constant. Note that $\tan(\pi n)$ always has + residue, unlike the usual S.W. $\sin(\pi n)$ which we correct there by using $\cot(-z)$. OK, now make sure that $f(n)$ has no (n) singularities, then unwind contour to B; make sure A. Circle goes away (it must). Then:

$$A \approx \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} f(n) \cot(\pi n) dn \approx \int_{-i\infty}^{i\infty} f(n+\frac{1}{2}) \tan(\pi n) dn$$

Next, move contour and go to a real integral:

$$A \approx \int_{-\infty}^{\infty} f(in+\frac{1}{2}) \tanh(\pi n) dn \approx 2 \int_0^{\infty} \text{Re} f(in+\frac{1}{2}) \tanh(\pi n) dn$$

Then do parts:

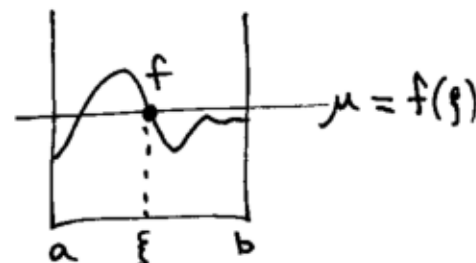
$$A \approx 2 \int_0^{\infty} \left[\int_{\text{ind}} \text{Re} f(in+\frac{1}{2}) \right] \text{sech}^2(\pi n) dn$$

And then approx to 13 places by

$$\approx 2 \int_0^5 F(n) \text{sech}^2(\pi n) dn.$$

Finally, do this by a "20 pt function" in gaussian quadrature!

Chapter 1



Introduction

1.1 The mean value theorems

First, we introduce the mean value for integrals. Assume that f is continuous, then

$$\mu = \frac{1}{b-a} \int_a^b f(t) dt = \frac{\int_a^b f(t) dt}{\int_a^b dt} . \quad \checkmark$$

This μ is called the arithmetic average or the mean value of f in the interval $[a, b]$. We state the mean value theorem for integrals:

Theorem 1.1 ([1, p.258] Mean value theorem for integrals).

If f is a continuous function on $[a, x]$, then there exists a number ξ_x in $[a, x]$ such that

$$\int_a^x f(t) dt = f(\xi_x)(x-a) . \quad (1.1)$$

This is a simple but very important mean value theorem of integral calculus.

! (It states that the mean value theorem of a continuous function in an interval belongs to the range of the function. And it asserts only the existence of at least

one ξ_x in the interval for which $f(\xi_x)$ is equal to the average value of f but gives no further information about the location ξ_x . In this thesis, we call ξ_x an integral mean point of f on $[a, x]$.

Instead of the simple arithmetic average we can form the weighted averages:

$$\mu = \frac{\int_a^x f(t)g(t)dt}{\int_a^x g(t)dt}$$

g is the weight function and $g(t) \geq 0$. We give the generalized mean value theorem for integrals:

Theorem 1.2 ([1, p.257] Generalized mean value theorem for integrals).

If f is a continuous function on $[a, x]$, g is integrable on $[a, x]$ and $g \geq 0$, then there exists a number ξ_x in $[a, x]$ such that

$$\int_a^x f(t)g(t)dt = f(\xi_x) \int_a^x g(t)dt . \quad (1.2)$$

Obviously, Theorem 1.1 is the special case when $g(t) = 1$.

Theorem 1.3 ([1, p.197] Classical mean value theorem).

Suppose f is continuous on the closed interval $[a, x]$ and differentiable on (a, x) , then for some c_x between $[a, x]$, we have

$$f'(c_x) = \frac{f(x) - f(a)}{x - a} . \quad (1.3)$$

We could estimate c_x as ξ_x in the same way as in (1.2). And we call c_x a differential mean point of f on $[a, x]$.

Theorem 1.4 ([1, p.198] Cauchy's mean value theorem).

See proof A.

Suppose f, g are continuous on $[a, x]$, differentiable on (a, x) . If $g'(t) \neq 0$ for any $t \in (a, x)$, then there exists $c_x \in (a, x)$ such that

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c_x)}{g'(c_x)} . \quad (1.4)$$

Here we call c_x a Cauchy differential mean point or simply Cauchy mean point.

Recently, there are a number of studies on the location of the integral mean point ξ_x , as $x \rightarrow a$. B. Zhang [7], improving B. Jacobson's result [4], showed that if f is r ($r \in N$) times differentiable at a , with $f'(a) = f''(a) = \dots = f^{(r-1)}(a) = 0$ but $f^{(r)}(a) \neq 0$, then

$$\lim_{x \rightarrow a} \frac{\xi_x - a}{x - a} = \frac{1}{(r+1)^{\frac{1}{r}}}.$$

Schwind-Ji-Koditschek [5] went on further. They allow f and g to have a 'singularity' at a . Namely they showed that if

$$\lim_{t \rightarrow a} \frac{f(t) - K}{(t - a)^r} = C_1$$

and

$$\lim_{t \rightarrow a} \frac{g(t)}{(t - a)^s} = C_2,$$

where $r \neq 0, s > -1, r + s > -1, C_1, C_2 \neq 0$, then

$$\lim_{x \rightarrow a} \frac{\xi_x - a}{x - a} = \left(\frac{s+1}{r+s+1} \right)^{\frac{1}{r}}.$$

On the other hand, Tong and Braza [6] studied the converse of the classical mean value theorem. They showed that if $f'(c)$ is not a local extremum, that c is a differential mean point.

It seems to be well-known that the four mean value theorems above are interrelated. In fact, Theorem 1.3 implies Theorem 1.1. Under some additional assumptions, Theorems 1.1 also implies Theorem 1.3. Similarly, Theorem 1.2 and Theorem 1.4 are roughly equivalent. Thus it is natural to study the estimates of the differential mean point and the converse of the mean value theorem for integrals.

In the next section, we shall discuss the relationships among the mean value theorems. In Chapter 2, we shall first discuss the works of B. Jacobson, B. Zhang and Schwind-Ji-Koditschek on the limiting position of the integral mean point ξ_x . Then we shall prove corresponding theorems for the differential mean point c_x of the classical mean value theorem and Cauchy mean value theorem.

In Chapter 3, we shall discuss the converse of the classical mean value theorem studied by Tong and Braza. Then, we prove a parallel theorem for the mean value theorem for integrals.

1.2 Relationships among the mean value theorems

It is known that Theorem 1.3 implies Theorem 1.1. There is a proof in the Calculus book written by Campbell and Dierker [3].

Proof of Theorem 1.1 ([3, p.209]).

Suppose $f : [a, x] \rightarrow R$ is continuous on $[a, x]$. Define $F(t) = \int_a^t f(s)ds$. Then F is continuous on $[a, x]$ and $F' = f$ on $[a, x]$. Hence by Theorem 1.3, there is some $c_x \in (a, x)$ such that

$$F(x) - F(a) = F'(c_x)(x - a) .$$

That means,

$$\int_a^x f(t)dt = f(c_x)(x - a) .$$

◇ ✓ (1.3) $\Rightarrow$ (1.1)

Conversely, suppose F is C^1 and $F' = f$. So

$$F(x) - F(a) = \int_a^x f(t)dt .$$

By Theorem 1.1 and above formula, we conclude that

$$F(x) - F(a) = \int_a^x f(t)dt = f(\xi_x)(x - a) .$$

Since F is a primitive of f , we could replace $f(\xi_x)$ by $F'(\xi_x)$ and obtain

$$F'(\xi_x) = \frac{F(x) - F(a)}{x - a} .$$

It is the same form as (1.4). Thus if F is C^1 , then Theorem 1.1 implies Theorem 1.3.

✓◇ (1.1 ⇒ 1.3)

Assume g is continuous and $g > 0$, then Theorem 1.4 implies Theorem 1.2.

The proof is as follows:

Define

$$F(t) = \int_a^t f(s)g(s)ds, \quad G(t) = \int_a^t g(s)ds .$$

F and G are continuous on $[a, x]$ and differentiable on (a, x) . Also

$$G(x) = \int_a^x g(t)dt > G(a) = 0 .$$

$g(t) \geq 0$.

Hence we apply Theorem 1.4 to obtain some $c_x \in (a, x)$ such that

$$F(x) - F(a) = \frac{F'(c_x)}{G'(c_x)}[G(x) - G(a)] .$$

Thus

$$\begin{aligned} \int_a^x f(t)g(t)dt &= \frac{f(c_x)g(c_x)}{g(c_x)} \int_a^x g(t)dt \quad \checkmark \\ &= f(c_x) \int_a^x g(t)dt. \quad \checkmark \quad = (1.2) \end{aligned}$$

◇✓ (1.4 $\Rightarrow$ 1.2)

Conversely, under additional assumption on f and g , then Theorem 1.2 implies Theorem 1.4. If f and g are both C^1 , let $\phi = \frac{f'}{g'}$ continuous, $\theta = g' > 0$ for simplicity, then by Theorem 1.2, there exists $c_x \in (a, x)$ such that

$$\int_a^x \phi(t)\theta(t)dt = \phi(c_x) \int_a^x \theta(t)dt .$$

Therefore,

$$\begin{array}{ccc} \downarrow & \downarrow & \downarrow \\ \int_a^x f'(t)dt = \frac{f'(c_x)}{g'(c_x)} \int_a^x g'(t)dt . \end{array}$$

So

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c_x)}{g'(c_x)} .$$

◇✓ (1.2 $\Rightarrow$ 1.4)

Math for the people, by the people.

[Encyclopedia](#) | [Requests](#) | [Forums](#) | [Docs](#) | [Random](#)

Login

[create new user](#)name: pass: [forget your password?](#)

Main Menu

[sections](#)[Encyclopædia](#)[Papers](#)[Books](#)[Expositions](#)[meta](#)[Requests](#) (71)[Orphanage](#) (10)[Unclass'd](#)[Unproven](#) (270)[Corrections](#) (120)[talkback](#)[Polls](#)[Forums](#)[Feedback](#)[Bug Reports](#)[downloads](#)[Snapshots](#)[PM Book](#)[information](#)[Docs](#)[Classification](#)[News](#)[Legalese](#)[History](#)[ChangeLog](#)[TODO List](#)

Cauchy
 proof of extended mean-value theorem

(Proof)

Let $f : [a, b] \rightarrow \mathbb{R}$ and $g : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b) . Define the function

$$h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)) - f(a)g(b) + f(b)g(a).$$

Because f and g are continuous on $[a, b]$ and differentiable on (a, b) , so is

h . Furthermore, $h(a) = h(b) = 0$, so by Rolle's theorem there exists a

$\xi \in (a, b)$ such that $h'(\xi) = 0$. This implies that

$$f'(\xi)(g(b) - g(a)) - g'(\xi)(f(b) - f(a)) = 0$$

and, if $g(b) \neq g(a)$,

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

"proof of extended mean-value theorem" is owned by [pbruin](#).

([view preamble](#))

View style: ☐ HTML with images ☒ 

See Also: [mean-value theorem](#)

Other names: [proof of Cauchy's mean-value theorem](#)

This object's [parent](#).

Cross-references: [Rolle's theorem](#), [function](#), [differentiable](#), [continuous](#)

This is [version 3](#) of [proof of extended mean-value theorem](#), born on 2002-11-12, modified 2003-01-27.

Object id is 3589, canonical name is ProofOfExtendedMeanValueTheorem.

Accessed 1690 times total.

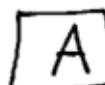
Classification:

AMS MSC: [26A06](#) (Real functions :: Functions of one variable :: One-variable calculus)

Pending Errata and Addenda

None.

[[View all 1](#)]





Math for the people, by the people.

[Encyclopedia](#) | [Requests](#) | [Forums](#) | [Docs](#) | [Random](#)

Login

[create new user](#)

name:

pass:

[forget your password?](#)

mean-value theorem

statement of

(Theorem)

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuous on the interval $[a, b]$ and differentiable on (a, b) . Then there exists a number $c : a < c < b$ such that

"classical mvt"

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \quad (1.3) \quad (1)$$

Main Menu

sections

[Encyclopedia](#)

[Papers](#)

[Books](#)

[Expositions](#)

meta

[Requests](#) (71)

[Orphanage](#) (10)

[Unclass'd](#)

[Unproven](#) (270)

[Corrections](#) (120)

talkback

[Polls](#)

[Forums](#)

[Feedback](#)

[Bug Reports](#)

downloads

[Snapshots](#)

[PM Book](#)

information

[Docs](#)

[Classification](#)

[News](#)

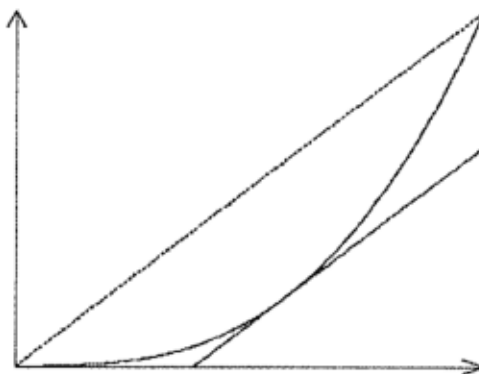
[Legalese](#)

[History](#)

[ChangeLog](#)

[TODO List](#)

The geometrical meaning of this theorem is illustrated in the picture:



The dashed line connects the points $(a, f(a))$ and $(b, f(b))$. There is c between a and b at which the tangent to f has the same slope as the dashed line.

The mean-value theorem is often used in the integral context: There is a $c \in [a, b]$ such that

"mvt for integrals"

$$(b - a)f(c) = \int_a^b f(x)dx. \quad (1.1) \quad (2)$$

"mean-value theorem" is owned by [mathwizrd](#). [[full author list](#) (2) | [owner history](#) (1)]

[\(view preamble\)](#)

View style: ☒ HTML with Images ☐ ☐

See Also: [Rolle's theorem](#), [intermediate value theorem](#), [extended mean-value theorem](#), [proof of extended mean-value theorem](#)

Attachments:

[proof of mean value theorem](#) (Proof) by saforres

[complex mean-value theorem](#) (Theorem) by matte

Cross-references: [integral](#), [tangent](#), [theorem](#), [differentiable](#), [interval](#), [continuous](#), [function](#)

There are 9 [references](#) to this object.

This is [version 6](#) of [mean-value theorem](#), born on 2002-02-15, modified 2004-07-21.

Object id is 1990, canonical name is MeanValueTheorem.

Accessed 3090 times total.

Classification:

AMS MSC: [26A06](#) (Real functions :: Functions of one variable :: One-variable calculus)



Math for the people, by the people.

[Encyclopedia](#) | [Requests](#) | [Forums](#) | [Docs](#) | [Random](#)

Login

▲ proof of mean value theorem

(Proof)

[create new user](#)

Define $h(x)$ on $[a, b]$ by

name:

pass:

[forget your password?](#)

$$h(x) = f(x) - f(a) - \left(\frac{f(b) - f(a)}{b - a} \right) (x - a)$$

Main Menu

Clearly, h is continuous on $[a, b]$, differentiable on (a, b) , and

[sections](#)

$$h(a) = f(a) - f(a) = 0$$

[Encyclopædia](#)

$$h(b) = f(b) - f(a) - \left(\frac{f(b) - f(a)}{b - a} \right) (b - a) = 0$$

[Papers](#)

[Books](#)

[Expositions](#)

Notice that h satisfies the conditions of Rolle's Theorem. Therefore, by Rolle's Theorem there exists $c \in (a, b)$ such that $h'(c) = 0$.

[meta](#)

[Requests](#) (71)

[Orphanage](#) (10)

[Unclass'd](#)

[Unproven](#) (270)

[Corrections](#) (120)

However, from the definition of h we obtain by differentiation that

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$$

[talkback](#)

Since $h'(c) = 0$, we therefore have

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

(1.3) classical

[Polls](#)

[Forums](#)

[Feedback](#)

[Bug Reports](#)

as required.

Rolle $\Rightarrow$ Classical

[downloads](#)

[Snapshots](#)

[PM Book](#)

Bibliography

[information](#)

[Docs](#)

[Classification](#)

[News](#)

[Legalese](#)

[History](#)

[ChangeLog](#)

[TODO List](#)

1

Michael Spivak, *Calculus*, 3rd ed., Publish or Perish Inc., 1994.

"proof of mean value theorem" is owned by [saforres](#).

([view preamble](#))

View style: ☐ HTML with images ☒

This object's parent.

Cross-references: [Rolle's theorem](#), [satisfies](#), [differentiable](#), [continuous](#)

This is version 2 of [proof of mean value theorem](#), born on 2002-05-28, modified 2002-05-29.
Object id is 2960, canonical name is ProofOfMeanValueTheorem.
Accessed 1190 times total.

Classification:

AMS MSC: [26A06](#) (Real functions :: Functions of one variable :: One-variable calculus)

Pending Errata and Addenda

None.

Discussion

Style: ☐ Threaded ☒ Expand: Order: ☐ Newest first ☒

No messages.



Math for the people, by the people.

[Encyclopedia](#) | [Requests](#) | [Forums](#) | [Docs](#) | [Random](#)

Login

[create new user](#)

name:

pass:

[forget your password?](#)

Main Menu

[sections](#)

[Encyclopædia](#)

[Papers](#)

[Books](#)

[Expositions](#)

[meta](#)

[Requests](#) (71)

[Orphanage](#) (10)

[Unclass'd](#)

[Unproven](#) (270)

[Corrections](#) (120)

[talkback](#)

[Polls](#)

[Forums](#)

[Feedback](#)

[Bug Reports](#)

[downloads](#)

[Snapshots](#)

[PM Book](#)

[information](#)

[Docs](#)

[Classification](#)

[News](#)

[Legalese](#)

[History](#)

[ChangeLog](#)

[TODO List](#)

proof of Rolle's theorem

(Proof)

Because f is continuous on a compact (closed and bounded) interval $I = [a, b]$, it attains its maximum and minimum values. In case $f(a) = f(b)$ is both the maximum and the minimum, then there is nothing more to say, for then f is a constant function and $f' \equiv 0$ on the whole interval I . So suppose otherwise, and f attains an extremum in the open interval (a, b) , and without loss of generality, let this extremum be a maximum, considering $-f$ in lieu of f as necessary. We claim that at this extremum $f(c)$ we have $f'(c) = 0$, with $a < c < b$.

To show this, note that $f(x) - f(c) \leq 0$ for all $x \in I$, because $f(c)$ is the maximum. By definition of the derivative, we have that

$$f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}.$$

Looking at the one-sided limits, we note that

$$R = \lim_{x \rightarrow c^+} \frac{f(x) - f(c)}{x - c} \leq 0$$

because the numerator in the limit is nonpositive in the interval I , yet $x - c > 0$, as x approaches c from the right. Similarly,

$$L = \lim_{x \rightarrow c^-} \frac{f(x) - f(c)}{x - c} \geq 0.$$

Since f is differentiable at c , the left and right limits must coincide, so $0 \leq L = R \leq 0$, that is to say, $f'(c) = 0$.

"proof of Rolle's theorem" is owned by [rmilson](#). [[full author list](#) (2) | [owner history](#) (1)]
(view preamble)

View style: ☒ HTML with images ☐

This object's parent.

Cross-references: [differentiable](#), [limit](#), [numerator](#), [one-sided limits](#), [derivative](#), [open interval](#), [extremum](#), [constant function](#), [interval](#), [bounded](#), [closed](#), [compact](#), [continuous](#)

This is version 2 of [proof of Rolle's theorem](#), born on 2002-05-27, modified 2004-02-26.
Object id is 2947, canonical name is ProofOfRolleTheorem.
Accessed 1319 times total.

Classification:

AMS MSC: [26A06](#) (Real functions :: Functions of one variable :: One-variable calculus)

Pending Errata and Addenda

None.

[[View all 2](#)]

