axiom™

The 30 Year Horizon

Manuel Bronstein *William Burge* *Timothy Daly*
James Davenport *Michael Dewar* *Martin Dunstan*
Albrecht Fortenbacher *Patrizia Gianni* *Johannes Grabmeier*
Jocelyn Guidry *Richard Jenks* *Larry Lambe*
Michael Monagan *Scott Morrison* *William Sit*
Jonathan Steinbach *Robert Sutor* *Barry Trager*
Stephen Watt *Jim Wen* *Clifton Williamson*

Volume 10: Axiom Algebra: Theory

Inclusion of names in the list of credits is based on historical information and is as accurate as possible. Inclusion of names does not in any way imply an endorsement but represents historical influence on Axiom development.

Cyril Alberga	Roy Adler	Richard Anderson
George Andrews	Henry Baker	Stephen Balzac
Yurij Baransky	David R. Barton	Gerald Baumgartner
Gilbert Baumslag	Fred Blair	Vladimir Bondarenko
Mark Botch	Alexandre Bouyer	Peter A. Broadbery
Martin Brock	Manuel Bronstein	Florian Bundschuh
William Burge	Quentin Carpent	Bob Caviness
Bruce Char	Cheekai Chin	David V. Chudnovsky
Gregory V. Chudnovsky	Josh Cohen	Christophe Conil
Don Coppersmith	George Corliss	Robert Corless
Gary Cornell	Meino Cramer	Claire Di Crescenzo
Timothy Daly Sr.	Timothy Daly Jr.	James H. Davenport
Jean Della Dora	Gabriel Dos Reis	Michael Dewar
Claire DiCrescendo	Sam Dooley	Lionel Ducos
Martin Dunstan	Brian Dupee	Dominique Duval
Robert Edwards	Heow Eide-Goodman	Lars Erickson
Richard Fateman	Bertfried Fauser	Stuart Feldman
Brian Ford	Albrecht Fortenbacher	George Frances
Constantine Frangos	Timothy Freeman	Korrinn Fu
Marc Gaetano	Rudiger Gebauer	Kathy Gerber
Patricia Gianni	Holger Gollan	Teresa Gomez-Diaz
Laureano Gonzalez-Vega	Stephen Gortler	Johannes Grabmeier
Matt Grayson	James Griesmer	Vladimir Grinberg
Oswald Gschnitzer	Jocelyn Guidry	Steve Hague
Vilya Harvey	Satoshi Hamaguchi	Martin Hassner
Ralf Hemmecke	Henderson	Antoine Hersen
Pietro Iglio	Richard Jenks	Kai Kaminski
Grant Keady	Tony Kennedy	Paul Kosinski
Klaus Kusche	Bernhard Kutzler	Larry Lambe
Frederic Lehobey	Michel Levaud	Howard Levy
Rudiger Loos	Michael Lucks	Richard Luczak
Camm Maguire	Bob McElrath	Michael McGettrick
Ian Meikle	David Mentre	Victor S. Miller
Gerard Milmeister	Mohammed Mobarak	H. Michael Moeller
Michael Monagan	Marc Moreno-Maza	Scott Morrison
Mark Murray	William Naylor	C. Andrew Neff
John Nelder	Godfrey Nolan	Arthur Norman
Jinzhong Niu	Michael O'Connor	Kostas Oikonomou
Julian A. Padget	Bill Page	Jaap Weel
Susan Pelzel	Michel Petitot	Didier Pinchon
Claude Quitte	Norman Ramsey	Michael Richardson
Renaud Rioboo	Jean Rivlin	Nicolas Robidoux
Simon Robinson	Michael Rothstein	Martin Rubey
Philip Santas	Alfred Scheerhorn	William Schelter
Gerhard Schneider	Martin Schoenert	Marshall Schor
Fritz Schwarz	Nick Simicich	William Sit
Elena Smirnova	Jonathan Steinbach	Christine Sundaresan
Robert Sutor	Moss E. Sweedler	Eugene Surowitz
James Thatcher	Baldir Thomas	Mike Thomas
Dylan Thurston	Barry Trager	Themos T. Tsikas
Gregory Vanuxem	Bernhard Wall	Stephen Watt
Juergen Weiss	M. Weller	Mark Wegman
James Wen	Thorsten Werther	Michael Wester
John M. Wiley	Berhard Will	Clifton J. Williamson
Stephen Wilson	Shmuel Winograd	Robert Wisbauer
Sandra Wityak	Waldemar Wiwianka	Knut Wolf

Contents

New Foreword

On October 1, 2001 Axiom was withdrawn from the market and ended life as a commercial product. On September 3, 2002 Axiom was released under the Modified BSD license, including this document. On August 27, 2003 Axiom was released as free and open source software available for download from the Free Software Foundation's website, Savannah.

Work on Axiom has had the generous support of the Center for Algorithms and Interactive Scientific Computation (CAISS) at City College of New York. Special thanks go to Dr. Gilbert Baumslag for his support of the long term goal.

The online version of this documentation is roughly 1000 pages. In order to make printed versions we've broken it up into three volumes. The first volume is tutorial in nature. The second volume is for programmers. The third volume is reference material. We've also added a fourth volume for developers. All of these changes represent an experiment in print-on-demand delivery of documentation. Time will tell whether the experiment succeeded.

Axiom has been in existence for over thirty years. It is estimated to contain about three hundred man-years of research and has, as of September 3, 2003, 143 people listed in the credits. All of these people have contributed directly or indirectly to making Axiom available. Axiom is being passed to the next generation. I'm looking forward to future milestones.

With that in mind I've introduced the theme of the "30 year horizon". We must invent the tools that support the Computational Mathematician working 30 years from now. How will research be done when every bit of mathematical knowledge is online and instantly available? What happens when we scale Axiom by a factor of 100, giving us 1.1 million domains? How can we integrate theory with code? How will we integrate theorems and proofs of the mathematics with space-time complexity proofs and running code? What visualization tools are needed? How do we support the conceptual structures and semantics of mathematics in effective ways? How do we support results from the sciences? How do we teach the next generation to be effective Computational Mathematicians?

The "30 year horizon" is much nearer than it appears.

Tim Daly
CAISS, City College of New York
November 10, 2003 ((iHy))

Chapter 1

Integration

An *elementary function* of a variable x is a function that can be obtained from the rational functions in x by repeatedly adjoining a finite number of nested logarithms, exponentials, and algebraic numbers or functions. Since $\sqrt{-1}$ is elementary, the trigonometric functions and their inverses are also elementary (when they are rewritten using complex exponentials and logarithms) as well as all the "usual" functions of calculus. For example,

$$\sin(x + \tan(x^3 - \sqrt{x^3 - x + 1})) \tag{1.1}$$

is elementary when rewritten as

$$\frac{\sqrt{-1}}{2}(e^{t-x\sqrt{-1}} - e^{x\sqrt{-1}-t}) \text{ where } t = \frac{1 - e^{2\sqrt{-1}(x^3 - \sqrt{x^3-x+1})}}{1 + e^{2\sqrt{-1}(x^3 - \sqrt{x^3-x+1})}}$$

This tutorial describes recent algorithmic solutions to the *problem of integration in finite terms*: to decide in a finite number of steps whether a given elementary funcction has an elementary indefinite integral, and to compute it explicitly if it exists. While this problem was studied extensively by Abel and Liouville during the last century, the difficulties posed by algebraic functions caused Hardy (1916) to state that "there is reason to suppose that no such method can be given". This conjecture was eventually disproved by Risch (1970), who described an algorithm for this problem in a series of reports [Ost1845, Ris68, Ris69a, Ris69b]. In the past 30 years, this procedure has been repeatedly improved, extended and refined, yielding practical algorithms that are now becoming standard and are implemented in most of the major computer algebra systems. In this tutorial, we outline the above algorithms for various classes of elementary functions, starting with rational functions and progressively increasing the class of functions up to general elementary functions. Proofs of correctness of the algorithms presented here can be found in several of the references, and are generally too long and too detailed to be described in this tutorial.

Notations: we write x for the variable of integration, and ' for the derivation d/dx. $\mathbb{Z}$,$\mathbb{Q}$,$\mathbb{R}$,and $\mathbb{C}$ denote respectively the integers, rational, real and complex numbers. All fields are commutative and, except when mentioned explicitly otherwise, have characteristic 0. If K is a field, then $\overline{K}$ denotes its algebraic closure. For a polynomial p, $\mathrm{pp}(p)$ denotes the primitive part of p, *i. e.* p divided by the gcd of its coefficients.

1.1 Rational Functions

By a *rational function*, we mean a quotient of polynomials in the integration variable x. This means that other functions can appear in the integrand, provided they do not involve x, hence that the coefficients of our polynomials in x lie in an arbitrary field K satisfying: $\forall a \in K,\ a^{'} = 0$.

1.1.1 The full partial-fraction algorithm

This method, which dates back to Newton, Leibniz, and Bernoulli, should not be used in practice, yet it remains the method found in most calculus tests and is often taught. Its major drawback is the factorization of the denominator of the integrand over the real or complex numbers. We outline it because it provides the theoretical foundations for all the subsequent algorithms. Let $f \in \mathbb{R}(x)$ be our integrand, and write $f = P + A/D$ where $P, A, D \in \mathbb{R}[x]$, $gcd(A, D) = 1$, and $deg(A) < deg(D)$. Let

$$D = c \prod_{i=1}^{n} (x - a_i)^{e_i} \prod_{j=1}^{m} (x^2 + b_j x + c_j)^{f_j}$$

be the irreducible factorization of D over $\mathbb{R}$, where c, the a_i's, b_j's and c_j's are in $\mathbb{R}$ and the e_i's and f_j's are positive integers. Computing the partial fraction decomposition of f, we get

$$f = P + \sum_{i=1}^{n} \sum_{k=1}^{e_i} \frac{A_{ik}}{(x - a_i)^k} + \sum_{j=1}^{m} \sum_{k=1}^{f_i} \frac{B_{jk} x + C_{jk}}{(x^2 + b_j x + c_j)^k}$$

where the A_{ik}'s, B_{jk}'s, and C_{jk}'s are in $\mathbb{R}$. Hence,

$$\int f = \int P + \sum_{i=1}^{n} \sum_{k=1}^{e_i} \int \frac{A_{ik}}{(x - a_i)^k} + \sum_{j=1}^{m} \sum_{k=1}^{f_i} \int \frac{B_{jk} x + C_{jk}}{(x^2 + b_j x + c_j)^k}$$

Computing $\int P$ poses no problem (it will for any other class of functions), and for the other terms we have

$$\int \frac{A_{ik}}{(x - a_i)^k} = \begin{cases} A_{ik}(x - a_i)^{1-k}/(1 - k) & \text{if } k > 1 \\ A_{i1} \log(x - a_i) & \text{if } k = 1 \end{cases} \tag{1.2}$$

and, noting that $b_j^2 - 4c_j < 0$ since $x^2 + b_j x + c_j$ is irreducible in $\mathbb{R}$[x].

$$\int \frac{B_{j1}x + C_{j1}}{(x^2 + b_j x + c_j)} = \frac{B_{j1}}{2}\log(x^2+b_jx+c_j)+\frac{2C_{j1} - b_j B_{j1}}{\sqrt{4c_j - b_j^2}} arctan\left(\frac{2x + b_j}{\sqrt{4c_j - b_j^2}}\right)$$

and for $k > 1$,

$$\begin{aligned}\int \frac{B_{jk}x + C_{jk}}{(x^2 + b_j x + c_j)^k} &= \frac{(2C_{jk} - b_j B_{jk})x + b_j C_{jk} - 2c_j B_{jk}}{(k-1)(4c_j - b_j^2)(x^2 + b_j x + c_j)^{k-1}} \\ &+ \int \frac{(2k-3)(2C_{jk} - b_j B_{jk})}{(k-1)(4c_j - b_j^2)(x^2 + b_j x + c_j)^{k-1}}\end{aligned}$$

This last formula is then used recursively until $k = 1$.

An alternative is to factor D linearly over $\mathbb{C}$: $D = \prod_{i=1}^{q}(x - \alpha_i)^{e_i}$, and then use (2) on each term of

$$f = P + \sum_{i=1}^{q}\sum_{j=1}^{e_i}\frac{A_{ij}}{(x - \alpha_i)^j} \tag{1.3}$$

Note that this alternative is applicable to coefficients in any field K, if we factor D linearly over its algebraic closure $\overline{K}$, and is equivalent to expanding f into its Laurent series at all its finite poles, since that series at $x = \alpha_i \in \overline{K}$ is

$$f = \frac{A_{ie_i}}{(x - \alpha_i)^{e_i}} + \cdots + \frac{A_{i2}}{(x - \alpha_i)^2} + \frac{A_{i1}}{(x - \alpha_i)} + \cdots$$

where the A_{ij}'s are the same as those in (3). Thus, this approach can be seen as expanding the integrand into series around all the poles (including ∞), then integrating the series termwise, and then interpolating for the answer, by summing all the polar terms, obtaining the integral of (3). In addition, this alternative shows that any rational function $f \in K(x)$ has an elementary integral of the form

$$\int f = v + c_1 \log(u_1) + \cdots + c_m \log(u_m) \tag{1.4}$$

where $v, u_1, \ldots, u_m \in \overline{K}(x)$ are the rational functions, and $c_1, \ldots, c_m \in \overline{K}$ are constants. The original Risch algorithm is essentially a generalization of this approach that searches for integrals of arbitrary elementary functions in a form similar to (4).

1.1.2 The Hermite reduction

The major computational inconvenience of the full partial fraction approach is the need to factor polynomials over $\mathbb{R}$, $\mathbb{C}$, or $\overline{K}$, thereby introducing algebraic numbers even if the integrand and its integral are both in $\mathbb{Q}(x)$. On the other hand, introducing algebraic numbers may be necessary, for example it is proven in [Ris69a] that any field containing an integral of $1/(x^2 + 2)$ must also contain $\sqrt{2}$. Modern research has yielded so-called "rational" algorithms that

- compute as much of the integral as possible with all calculations being done in $K(x)$, and
- compute the minimal algebraic extension of K necessary to express the integral

The first rational algorithms for integration date back to the 19^{th} century, when both Hermite[Her1872] and Ostrogradsky[Ost1845] invented methods for computing the v of (4) entirely within $K(x)$. We describe here only Hermite's method, since it is the one that has been generalized to arbitrary elementary functions. The basic idea is that if an irreducible $p \in K[x]$ appears with multiplicity $k > 1$ in the factorization of the denominator of the integrand, then (2) implies that it appears with multiplicity $k - 1$ in the denominator of the integral. Furthermore, it is possible to compute the product of all such irreducibles for each k without factoring the denominator into irreducibles by computing its *squarefree factorization*, *i.e* a factorization $D = D_1 D_2^2 \cdots D_m^m$, where each D_i is squarefree and $gcd(D_i, D_j) = 1$ for $i \neq j$. A straightforward way to compute it is as follows: let $R = gcd(D, D^{'})$, then $R = D_2 D_2^3 \cdots D_m^{m-1}$, so $D/R = D_1 D_2 \cdots D_m$ and $gcd(R, D/R) = D_2 \cdots D_m$, which implies finally that

$$D_1 = \frac{D/R}{gcd(R, D/R)}$$

Computing recursively a squarefree factorization of R completes the one for D. Note that [Yu76] presents a more efficient method for this decomposition. Let now $f \in K(x)$ be our integrand, and write $f = P + A/D$ where $P, A, D \in K[x]$, $gcd(A, D) = 1$, and $deg(A) < deg(D)$. Let $D = D_1 D_2^2 \cdots D_m^m$ be a squarefree factorization of D and suppose that $m \geq 2$ (otherwise D is already squarefree). Let then $V = D_m$ and $U = D/V^m$. Since $gcd(UV^{'}, V) = 1$, we can use the extended Euclidean algorithm to find $B, C \in K[x]$ such that

$$\frac{A}{1-m} = BUV^{'} + CV$$

and $deg(B) < deg(V)$. Multiplying both sides by $(1-m)/(UV^m)$ gives

$$\frac{A}{UV^m} = \frac{(1-m)BV^{'}}{V^m} + \frac{(1-m)C}{UV^{m-1}}$$

so, adding and subtracting $B^{'}/V^{m-1}$ to the right hand side, we get

$$\frac{A}{UV^m} = \left(\frac{B^{'}}{V^{m-1}} - \frac{(m-1)BV^{'}}{V^m}\right) + \frac{(1-m)C - UB^{'}}{UV^{m-1}}$$

and integrating both sides yields

$$\int \frac{A}{UV^m} = \frac{B}{V^{m-1}} + \int \frac{(1-m)C - UB^{'}}{UV^{m-1}}$$

so the integrand is reduced to one with a smaller power of V in the denominator. This process is repeated until the denominator is squarefree, yielding $g, h \in K(x)$ such that $f = g^{'} + h$ and h has a squarefree denominator.

1.1.3 The Rothstein-Trager and Lazard-Rioboo-Trager algorithms

Following the Hermite reduction, we only have to integrate fractions of the form $f = A/D$ with $deg(A) < deg(D)$ and D squarefree. It follows from (2) that

$$\int f = \sum_{i=1}^{n} a_i \log(x - \alpha_i)$$

where the α_i's are the zeros of D in $\overline{K}$, and the a_i's are the residues of f at the α_i's. The problem is then to compute those residues without splitting D. Rothstein [Ro77] and Trager [Tr76] independently proved that the α_i's are exactly the zeros of

$$R = resultant_x(D, A - tD^{'}) \in K[t] \tag{1.5}$$

and that the splitting field of R over K is indeed the minimal algebraic extension of K necessary to express the integral in the form (4). The integral is then given by

$$\int \frac{A}{D} = \sum_{i=1}^{m} \sum_{a|R_i(a)=0} a \log(\gcd(D, A - aD^{'})) \tag{1.6}$$

where $R = \prod_{i=1}^{m} R_i^{e_i}$ is the irreducible factorization of R over K. Note that this algorithm requires factoring R into irreducibles over K, and computing greatest common divisors in $(K[t]/(R_i))[x]$, hence computing with algebraic numbers. Trager and Lazard & Rioboo [LR90] independently discovered that those computations can be avoided, if one uses the subresultant PRS algorithm to compute the resultant of (5): let $(R_0, R_1, \ldots R_k \neq 0, 0, \ldots)$ be the subresultant PRS with respect to x of D and $A - tD^{'}$ and $R = Q_1 Q_2^2 \ldots Q_m^m$ be a *squarefree* factorization of their resultant. Then,

$$\sum_{a|Q_i(a)=0} a \log(\gcd(D, A - aD^{'})) =$$

$$\begin{cases} \sum_{a|Q_i(a)=0} a \log(D) & \text{if } i = deg(D) \\ \sum_{a|Q_i(a)=0} a \log(\mathrm{pp}_x(R_{k_i})(a, x)) & \text{where } deg(R_{k_i}) = i, 1 \leq k_i \leq n \\ & \text{if } i < deg(D) \end{cases}$$

Evaluating $\mathrm{pp}_x(R_{k_i})$ at $t = a$ where a is a root of Q_i is equivalent to reducing each coefficient with respect to x of $\mathrm{pp}_x(R_{k_i})$ module Q_i, hence computing in the algebraic extension $K[t]/(Q_i)$. Even this step can be avoided: it is in fact sufficient to ensure that Q_i and the leading coefficient with respect to x of R_{k_i} do not have a nontrivial common factor, which implies then that the remainder by Q_i is nonzero, see [Mul97] for details and other alternatives for computing $\mathrm{pp}_x(R_{k_i})(a, x)$

1.2 Algebraic Functions

By an *algebraic function*, we mean an element of a finitely generated algebraic extension E of the rational function field $K(x)$. This includes nested radicals and implicit algebraic functions, not all of which can be expressed by radicals. It turns out that the algorithms we used for rational functions can be extended to algebraic functions, but with several difficulties, the first one being to define the proper analogues of polynomials, numerators and denominators. Since E is algebraic over $K(x)$, for any $\alpha \in E$, there exists a polynomial $p \in K[x][y]$ such that $p(x,\alpha) = 0$. We say that $\alpha \in E$ is *integral over* $K[x]$ if there is a polynomial $p \in K[x][y]$, *monic in* y, such that $p(x,\alpha) = 0$. Integral elements are analogous to polynomials in that their value is defined for any $x \in \overline{K}$ (unlike non-integral elements, which must have at least one pole in $\overline{K}$). The set

$$\mathbf{O}_{K[x]} = \{\alpha \in E \text{ such that } \alpha \text{ is integral over } K[x]\}$$

is called the *integral closure of* $K[x]$ *in* E. It is a ring and a finitely generated $K[x]$-module. Let $\alpha \in E^*$ be any element and $p = \sum_{i=0}^m a_i y^i \in K[x][y]$ be such that $p(x,\alpha) = 0$ and $a_m \neq 0$. Then, $q(x, a_m y) = 0$ where $q = y^m + \sum_{i=0}^{m-1} a_i a_m^{m-i-1} y^i$ is monic in y, so $a_m y \in \mathbf{O}_{K[x]}$. We need a canonical representation for algebraic functions similar to quotients of polynomials for rational functions. Expressions as quotients of integral functions are not unique, for example, $\sqrt{x}/x = x/\sqrt{x}$. Hoever, E is a finite-dimensional vector space over $K(x)$, so let $n = [E : K(x)]$ and $w = (w_1, \ldots, w_n)$ be any basis for E over $K(x)$. By the above remark, there are $a_1, \ldots, a_n \in K(x)^*$ such that $a_i w_i \in \mathbf{O}_{K[x]}$ for each i. Since $(a_1 w_1, \ldots, a_n w_n)$ is also a basis for E over $K(x)$, we can assume without loss of generality that the basis w is composed of integral elements. Any $\alpha \in E$ can be written uniquely as $\alpha = \sum_{i=1}^n f_i w_i$ for $f_1, \ldots, f_n \in K(x)$, and putting the f_i's over a monic common denominator $D \in K[x]$, we get an expression

$$\alpha = \frac{A_1 w_1 + \ldots + A_n w_n}{D}$$

where $A_1, \ldots, A_n \in K[x]$ and $\gcd(D, A_1, \ldots, A_n) = 1$. We call $\sum_{i=1}^n A_i w_i \in \mathbf{O}_{K[x]}$ and $D \in K[x]$ respectively the *numerator* and *denominator* of α with respect to w. They are defined uniquely once the basis w is fixed.

1.2.1 The Hermite reduction

Now that we have numerators and denominators for algebraic functions, we can attempt to generalize the Hermite reduction of the previous section, so let $f \in E$ be our integrand, $w = (w_1, \ldots, w_n) \in {\mathbf{O}_{K[n]}}^n$ be a basis for E over $K(x)$ and let $\sum_{i=1}^m A_i w_i \in \mathbf{O}_{K[x]}$ and $D \in K[x]$ be the numerator and denominator of f with respect to w, Let $D = D_1 D_2^2 \ldots D_m^m$ be a squarefree factorization of D and suppose that $m \geq 2$. Let then $V = D_m$ and $U = D/V^m$, and we ask whether we can compute $B = \sum_{i=1}^n B_i w_i \in \mathbf{O}_{K[x]}$ and $h \in E$ such that $deg(B_i) < deg(V)$

for each i,

$$\int \frac{\sum_{i=1}^{n} A_i w_i}{UV^m} = \frac{B}{V^{m-1}} + \int h \tag{1.7}$$

and the denominator of h with respect to w has no factor of order m or higher. This turns out to reduce to solving the following linear system

$$f_1 S_1 + \ldots + f_n S_n = A_1 w_1 + \ldots + A_n w_n \tag{1.8}$$

for $f_1, \ldots, f_n \in K(x)$, where

$$S_i = UV^m \left(\frac{w_i}{V^{m-1}}\right)' \quad \text{for } 1 \le i \le n \tag{1.9}$$

Indeed, suppose that (8) has a solution $f_1, \ldots, f_n \in K(x)$, and write $f_i = T_i/Q$, where $Q, T_1, \ldots, T_n \in K[x]$ and $\gcd(Q, T_1, \ldots, T_n) = 1$. Suppose further that $\gcd(Q, V) = 1$. Then, we can use the extended Euclidean algorithm to find $A, R \in K[x]$ such that $AV + RQ = 1$, and Euclidean division to find $Q_i, B_i \in K[x]$ such that $deg(B_i) < deg(V)$ when $B_i \neq 0$ and $RT_i = VQ_i + B_i$ for each i. We then have

$$\begin{aligned}
h &= f - \left(\frac{\sum_{i=1}^{n} B_i w_i}{V^{m-1}}\right)' \\
&= \frac{\sum_{i=1}^{n} A_i w_i}{UV^m} - \frac{\sum_{i=1}^{n} B_i' w_i}{V^{m-1}} - \sum_{i=1}^{n} (RT_i - VQ_i)\left(\frac{w_i}{V^{m-1}}\right)' \\
&= \frac{\sum_{i=1}^{n} A_i w_i}{UV^m} - \frac{R\sum_{i=1}^{n} T_i S_i}{UV^m} + V\sum_{i=1}^{n} Q_i \left(\frac{w_i}{V^{m-1}}\right)' - \frac{\sum_{i=1}^{n} B_i' w_i}{V^{m-1}} \\
&= \frac{(1-RQ)\sum_{i=1}^{n} A_i w_i}{UV^m} + \frac{\sum_{i=1}^{n} Q_i w_i'}{V^{m-2}} - (m-1)V' \frac{\sum_{i=1}^{n} Q_i w_i}{V^{m-1}} - \frac{\sum_{i=1}^{n} B_i' w_i}{V^{m-1}} \\
&= \frac{\sum_{i=1}^{n} AA_i w_i}{UV^{m-1}} - \frac{\sum_{i=1}^{n}((m-1)V'Q_i + B_i')w_i}{V^{m-1}} + \frac{\sum_{i=1}^{n} Q_i w_i'}{V^{m-2}}
\end{aligned}$$

Hence, if in addition the denominator of h has no factor of order m or higher, then $B = \sum_{i=1}^{n} B_i w_i \in \mathbf{O}_{K[x]}$ and h solve (7) and we have reduced the integrand. Unfortunately, it can happen that the denominator of h has a factor of order m or higher, or that (8) has no solution in $K(x)$ whose denominator is coprime with V, as the following example shows.

Example 1 Let $E = K(x)[y]/(y^4 + (x^2 + x)y - x^2)$ *with basis* $w = (1, y, y^2, y^3)$ *over* $K(x)$ *and consider the integrand*

$$f = \frac{y^3}{x^2} = \frac{w_4}{x^2} \in E$$

We have $D = x^2$, so $U = 1, V = x$ and $m = 2$. Then, $S_1 = x^2(1/x)^{'} = -1$,

$$\begin{aligned}
S_2 &= x^2 \left(\frac{y}{x}\right)^{'} \\
&= \frac{24(1-x^2)y^3+32x(1-x)y^2-(9x^4+45x^3+209x^2+63x+18)y-18x(x^3+x^2-x-1)}{27x^4+108x^3+418x^2+108x+27} \\
S_3 &= x^2 \left(\frac{y^2}{x}\right)^{'} \\
&= \frac{64x(1-x)y^3+9(x^4+2x^3-2x-1)y^2+12x(x^3+x^2-x-1)y+48x^2(1-x^2)}{27x^4+108x^3+418x^2+108x+27}
\end{aligned}$$

and

$$\begin{aligned}
S_4 &= x^2 \left(\frac{y^3}{x}\right)^{'} \\
&= \frac{(27x^4+81x^3+209x^2+27x)y^3+18x(x^3+x^2-x-1)y^2+24x^2(x^2-1)y+96x^3(1-x)}{27x^4+108x^3+418x^2+108x+27}
\end{aligned}$$

so (8) becomes

$$M \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{1.10}$$

where

$$M = \begin{pmatrix} -1 & \frac{-18x(x^3+x^2-x-1)}{F} & \frac{48x^2(1-x^2)}{F} & \frac{96x^3(1-x)}{F} \\ 0 & \frac{-(9x^4+45x^3+209x^2+63x+18)}{F} & \frac{12x(x^3+x^2-x-1)}{F} & \frac{24x^2(x^2-1)}{F} \\ 0 & \frac{32x(1-x)}{F} & \frac{9(x^4+2x^3-2x-1)}{F} & \frac{18x(x^3+x^2-x-1)}{F} \\ 0 & \frac{24(1-x^2)}{F} & \frac{64x(1-x)}{F} & \frac{(27x^4+81x^3+209x^2+27x)}{F} \end{pmatrix}$$

and $F = 27x^4 + 108x^3 + 418x^2 + 108x + 27$. The system (10) admits a unique solution $f_1 = f_2 = 0, f_3 = -2$ and $f_4 = (x + 1)/x$, whose denominator is not coprime with V, so the Hermite reduction is not applicable.

The above problem was first solved by Trager [Tr84], who proved that if w is an *integral basis, i.e.* its elements generate $\mathbf{O}_{K[x]}$ over $K[x]$, then the system (8) always has a unique solution in $K(x)$ when $m > 1$, and that solution always has a denominator coprime with V. Furthermore, the denominator of each $w_i^{'}$ must be squarefree, implying that the denominator of h is a factor of FUV^{m-1} where $F \in K[x]$ is squarefree and coprime with UV. He also described an algorithm for computing an integral basis, a necessary preprocessing for his Hermite reduction. The main problem with that approach is that computing the integral basis, whether by the method of [Tr84] or the local alternative [vH94], can be in general more expansive than the rest of the reduction process. We describe here the lazy Hermite reduction [Bro98], which avoids the precomputation of an integral basis. It is based on the observation that if $m > 1$ and (8) does not have a solution allowing us to perform the reduction, then either

- the S_i's are linearly dependent over $K(x)$, or
- (8) has a unique solution in $K(x)$ whose denominator has a nontrivial common factor with V, or
- the denominator of some w_i is not squarefree

In all of the above cases, we can replace our basis w by a new one, also made up of integral elements, so that that $K[x]$-module generated by the new basis strictly contains the one generated by w:

Theorem 1 ([Bro98]) *Suppose that $m \geq 2$ and that $\{S_1, \ldots, S_n\}$ as given by (9) are linearly dependent over $K(x)$, and let $T_1, \ldots, T_n \in K[x]$ be not all 0 and such that $\sum_{i=1}^n T_i S_i = 0$. Then,*

$$w_0 = \frac{U}{V} \sum_{i=1}^{n} T_i w_i \in \mathbf{O}_{K[x]}$$

Furthermore, if $\gcd(T_1, \ldots, T_n) = 1$ then $w_0 \notin K[x]w_1 + \cdots + K[x]w_n$.

Theorem 2 ([Bro98]) *Suppose that $m \geq 2$ and that $\{S_1, \ldots, S_n\}$ as given by (9) are linearly independent over $K(x)$, and let $Q, T_1, \ldots, T_n \in K[x]$ be such that*

$$\sum_{i=1}^{n} A_i w_i = \frac{1}{Q} \sum_{i=1}^{n} T_i S_i$$

Then,

$$w_0 = \frac{U(V/\gcd(V,Q))}{\gcd(V,Q)} \sum_{i=1}^{n} T_i w_i \in \mathbf{O}_{K[x]}$$

Furthermore, if $\gcd(Q, T_1, \ldots, T_n) = 1$ and $\deg(\gcd(V,Q)) \geq 1$, then $w_0 \notin K[x]w_1 + \cdots + K[x]w_n$.

Theorem 3 ([Bro98]) *Suppose that the denominator F of some w_i is not squarefree, and let $F = F_1 F_2^2 \cdots F_k^k$ be its squarefree factorization. Then,*

$$w_0 = F_1 \cdots F_k w_i' \in \mathbf{O}_{K[x]} \backslash (K[x]w_1 + \cdots + K[x]w_n).$$

The lazy Hermite reduction proceeds by solving the system (8) in $K(x)$. Either the reduction will succeed, or one of the above theorems produces an element $w_0 \in \mathbf{O}_{K[x]} \backslash (K[x]w_1 + \cdots + K[x]w_n)$. Let then $\sum_{i=1}^n C_i w_i$ and F be the numerator and denominator of w_0 with respect to w. Using Hermitian row reduction, we can zero out the last row of

$$\begin{pmatrix} F & & & \\ & F & & \\ & & \ddots & \\ & & & F \\ C_1 & C_2 & \cdots & C_n \end{pmatrix}$$

obtaining a matrix of the form

$$\begin{pmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,n} \\ C_{2,1} & C_{2,2} & \cdots & C_{2,n} \\ \vdots & \vdots & & \vdots \\ C_{n,1} & C_{n,2} & \cdots & C_{n,n} \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

with $C_{ij} \in K[x]$. Let $\overline{w}_i = (\sum_{j=1}^n C_{ij} w_j)/F$ for $1 \le i \le n$. Then, $\overline{w} = (\overline{w}_1, \ldots, \overline{w}_n)$ is a basis for E over K and

$$K[x]\overline{w}_1 + \cdots + K[x]\overline{w}_n = K[x]w_1 + \cdots + K[x]w_n + K[x]w_0$$

is a submodule of $\mathbf{O}_{K[x]}$, which strictly contains $K[x]w_1 + \cdots + K[x]w_n$, since it contains w_0. Any strictly increasing chain of submodules of $\mathbf{O}_{K[x]}$ must stabilize after a finite number of steps, which means that this process produces a basis for which either the Hermite reduction can be carried out, or for which f has a squarefree denominator.

Example 2 Continuing example 1 for which the Hermite reduction failed, Theorem 2 implies that

$$w_0 = \frac{1}{x}(-2xw_3 + (x+1)w_4) = (-2xy^2 + (x+1)y^3)x \in \mathbf{O}_{K[x]}$$

Performing a Hermitian row reduction on

$$\begin{pmatrix} x & & & \\ & x & & \\ & & x & \\ & & & x \\ 0 & 0 & -2x & x+1 \end{pmatrix}$$

yields

$$\begin{pmatrix} x & & & \\ & x & & \\ & & x & \\ & & & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

so the new basis is $\overline{w} = (1, y, y^2, y^3/x)$, and the denominator of f with respect to $\overline{w}$ is 1, which is squarefree.

1.2.2 Simple radical extensions

The integration algorithm becomes easier when E is a simple radical extension of $K(x)$, *i.e.* $E = K(x)[y]/(y^n - a)$ for some $a \in K(x)$. Write $a = A/D$ where $A, D \in K[x]$, and let $AD^{n-1} = A_1 A_2^2 \cdots A_k^k$ be a squarefree factorization of AD^{n-1}. Writing $i = nq_i + r_i$, for $1 \le i \le k$, where $0 \le r_i < n$, let $F = A_1^{q_1} \cdots A_k^{q_k}$, $H = A_1^{r_1} \cdots A_k^{r_k}$ and $z = yD/F$. Then,

$$z^n = \left(y\frac{D}{F}\right)^n = \frac{y^n D^n}{F^n} = \frac{AD^{n-1}}{F} = A_1^{r_1} \cdots A_k^{r_k} = H$$

Since $r_i < n$ for each i, the squarefree factorization of H is of the form $H = H_1 H_2^2 \cdots H_m^m$ with $m < n$. An integral basis is then $w = (w_1, \ldots, w_n)$ where

$$w_i = \frac{z^{i-1}}{\prod_{j=1}^m H_j^{\lfloor (i-1)j/n \rfloor}} \quad 1 \le i \le n \tag{1.11}$$

and the Hermite reductions with respect to the above basis is always guaranteed to succeed. Furthermore, when using that basis, the system (8) becomes diagonal and its solution can be written explicitly: writing $D_i = \prod_{j=1}^m H_j^{\lfloor ij/n \rfloor}$ we have

$$\begin{aligned} S_i &= UV^m \left(\frac{w_i}{V^{m-1}}\right)' = UV^m \left(\frac{z^{i-1}}{D_{i-1}V^{m-1}}\right)' \\ &= UV^m \left(\frac{i-1}{n}\frac{H'}{H} - \frac{D_{i-1}{}'}{D_{i-1}} - (m-1)\frac{V'}{V}\right)\left(\frac{z^{i-1}}{D_{i-1}V^{m-1}}\right) \\ &= U\left(V\left(\frac{i-1}{n}\frac{H'}{H} - \frac{D_{i-1}{}'}{D_{i-1}}\right) - (m-1)V'\right) w_i \end{aligned}$$

so the unique solution of (8) in $K(x)$ is

$$f_i = \frac{A_i}{U\left(V\left(\frac{i-1}{n}\frac{H'}{H} - \frac{D_{i-1}{}'}{D_{i-1}}\right) - (m-1)V'\right)} \quad \text{for } 1 \le i \le n \tag{1.12}$$

and it can be shown that the denominator of each f_i is coprime with V when $m \ge 2$.

Example 3 *Consider*

$$\int \frac{(2x^8+1)\sqrt{(x^8+1)}}{x^{17}+2x^9+x}\, dx$$

The integrand is

$$f = \frac{(2x^8+1)y}{x^{17}+2x^9+x} \in E = \mathbb{Q}(x)[y]/(y^2 - x^8 - 1)$$

so $H = x^8 + 1$ which is squarefree, implying that the integral basis (11) is $(w_1, w_2) = (1, y)$. The squarefree factorization of $x^{17} + 2x^9 + x$ is $x(x^8+1)^2$ so $U = x$, $V = x^8 + 1$, $m = 2$, and the solution (12) of (8) is

$$f_1 = 0, \quad f_2 = \frac{2x^8+1}{x\left((x^8+1)\frac{1}{2}\frac{8x^7}{x^8+1} - 8x^7\right)} = -\frac{(2x^8+1)/4}{x^8}$$

We have $Q = x^8$, so $V - Q = 1$, $A = 1$, $R = -1$ and $RQf_2 = V/2 - 1/4$, implying that

$$B = -\frac{y}{4} \quad \text{and} \quad h = f - \left(\frac{B}{V}\right)' = \frac{y}{x(x^8+1)}$$

solve (7), i.e.

$$\int \frac{(2x^8+1)\sqrt{(x^8+1)}}{x^{17}+2x^9+x}\,dx = -\frac{\sqrt{x^8+1}}{4(x^8+1)} + \int \frac{\sqrt{x^8+1}}{x(x^8+1)}\,dx$$

and the remaining integrand has a squarefree denominator.

1.2.3 Liouville's Theorem

Up to this point, the algorithms we have presented never fail, yet it can happen that an algebraic function does not have an elementary integral, for example

$$\int \frac{x\ dx}{\sqrt{1-x^3}}$$

which is not an elementary function of x. So we need a way to recognize such functions before completing the integration algorithm. Liouville was the first to state and prove a precise theorem from Laplace's observation that we can restrict the elementary integration problem by allowing only new logarithms to appear linearly in the integral, all the other terms appearing in the integral being already in the integrand.

Theorem 4 (Liouville [Lio1833a, Lio1833b]) *Let E be an algebraic extension of the rational function field $K(x)$, and $f \in E$. If f has an elementary integral, then there exist $v \in E$, constants $c_1, \ldots, c_n \in \overline{K}$ and $u_1, \ldots, u_k \in E(c_1, \ldots, c_k)^*$ such that*

$$f = v' + c_1\frac{u_1'}{u_1} + \cdots + c_k\frac{u_k'}{u_k} \tag{1.13}$$

The above is a restriction to algebraic functions of the strong Liouville Theorem, whose proof can be found in [Bro97, Ris69b]. An elegant and elementary algebraic proof of a slightly weaker version can be found in [Ro72]. As a consequence, we can look for an integral of the form (4), Liouville's Theorem guaranteeing that there is no elementary integral if we cannot find one in that form. Note that the above theorem does not say that every integral must have the above form, and in fact that form is not always the most convenient one, for example,

$$\int \frac{dx}{1+x^2} = arctan(x) = \frac{\sqrt{-1}}{2}\log\left(\frac{\sqrt{-1}+x}{\sqrt{-1}-x}\right)$$

1.2.4 The integral part

Following the Hermite reduction, we can assume that we have a basis $w = (w_1, \dots, w_n)$ of E over $K(x)$ made of integral elements such that our integrand is of the form $f = \sum_{i=1}^{n} A_i w_i / D$ where $D \in K[x]$ is squarefree. Given Liouville's Theorem, we now have to solve equation (13) for v, $u_1, \dots, u_k$ and the constants $c_1, \dots, c_k$. Since D is squarefree, it can be shown that $v \in \mathbf{O}_{K[x]}$ for any solution, and in fact v corresponds to the polynomial part of the integral of rational functions. It is however more difficult to compute than the integral of polynomials, so Trager [Tr84] gave a change of variable that guarantees that either $v' = 0$ or f has no elementary integral. In order to describe it, we need to define the analogue for algebraic functions of having a nontrivial polynomial part: we say that $\alpha \in E$ is *integral at infinity* if there is a polynomial $p = \sum_{i=1}^{m} a_i y^i \in K[x][y]$ such that $p(x, \alpha) = 0$ and $deg(a_m) \geq deg(a_i)$ for each i. Note that a rational function $A/D \in K(x)$ is integral at infinity if and only if $deg(A) \leq deg(D)$ since it is a zero of $Dy - A$. When $\alpha - E$ is not integral at infinity, we say that it has a *pole at infinity*. Let

$$\mathbf{O}_\infty = \{\alpha \in E \text{ such that } \alpha \text{ is integral at infinity}\}$$

A set $(b_1, \dots, b_n) \in E^n$ is called *normal at infinity* if there are $r_1, \dots, r_n \in K(x)$ such that every $\alpha \in \mathbf{O}_\infty$ can be written as $\alpha = \sum_{i=1}^{n} B_i r_i b_i / C$ where $C, B_1, \dots, B_n \in K[x]$ and $deg(C) \geq deg(B_i)$ for each i. We say that the differential $\alpha\ dx$ is integral at infinity if $\alpha x^{1+1/r} \in \mathbf{O}_\infty$ where r is the smallest ramification index at infinity. Trager [Tr84] described an algorithm that converts an arbitrary integral basis $w_1, \dots, w_n$ into one that is also normal at infinity, so the first part of his integration algorithm is as follows:

1. Pick any basis $b = (b_1, \dots, b_n)$ of E over $K(x)$ that is composed of integral elements.

2. Pick an integer $N \in \mathbb{Z}$ that is not zero of the denominator of f with respect to b, nor of the discriminant of E over $K(x)$, and perform the change of variable $x = N + 1/z$, $dx = -dz/z^2$ on the integrand.

3. Compute an integral basis w for E over $K(z)$ and make it normal at infinity

4. Perform the Hermite reduction on f using w, this yields $g, h \in E$ such that $\int f\ dz = g + \int h\ dz$ and h has a squarefree denominator with respect to w.

5. If hz^2 has a pole at infinity, then $\int f\ dz$ and $\int h\ dz$ are not elementary functions

6. Otherwise, $\int h\ dz$ is elementary if and only if there are constants $c_1, \dots, c_k \in \overline{K}$ and $u_1, \dots, u_k \in E(c_1, \dots, c_k)^*$ such that

$$h = \frac{c_1}{u_1}\frac{du_1}{dz} + \cdots + \frac{c_k}{u_k}\frac{du_k}{dz} \tag{1.14}$$

The condition that N is not a zero of the denominator of f with respect to b implies that the $f\ dz$ is integral at infinity after the change of variable, and Trager proved that if $h\ dz$ is not integral at infinity after the Hermite reduction, then $\int h\ dz$ and $\int f\ dz$ are not elementary functions. The condition that N is not a zero of the discriminant of E over $K(x)$ implies that the ramification indices at infinity are all equal to 1 after the change of variable, hence that $h\ dz$ is integral at infinity if and only if $hz^2 \in \mathbf{O}_\infty$. That second condition on N can be disregarded, in which case we must replace hz^2 in step 5 by $hz^{1+1/r}$ where r is the smallest ramification index at infinity. Note that $hz^2 \in \mathbf{O}_\infty$ implies that $hz^{1+1/r} \in \mathbf{O}_\infty$, but not conversely. Finally, we remark that for simple radical extensions, the integral basis (11) is already normal at infinity.

Alternatively, we can use lazy Hermite reduction in the above algorithm: in step 3, we pick any basis made of integral elements, then perform the lazy Hermite reduction in step 4. If $h \in K(z)$ after the Hermite reduction, then we can complete the integral without computing an integral basis. Otherwise, we compute an integral basis and make it normal at infinity between steps 4 and 5. This lazy variant can compute $\int f\ dx$ whenever it is an element of E without computing an integral basis.

1.2.5 The logarithmic part

Following the previous sections, we are left with solving equation (14) for the constants $c_1, \ldots, c_k$ and for $u_1, \ldots, u_k$. We must make at this point the following additional assumptions:

- we have an integral primitive element for E over $K(z)$, *i.e.* $y \in \mathbf{O}_{K[z]}$ such that $E = K(z)(y)$,
- $[E : K(z)] = [E : \overline{K}(z)]$, *i.e.* the minimal polynomial for y over $K[z]$ is absolutely reducible, and
- we have an integral basis $w = (w_1, \ldots, w_n)$ for E over $K(z)$, and w is normal at infinity

A primitive element can be computed by considering linear combinations of the generators of E over $K(x)$ with random coefficients in $K(x)$, and Trager [Tr84] describes an absolute factorization algorithm, so the above assumptions can be ensured, although those steps can be computationally very expensive, except in the case of simple radical extensions. Before describing the second part of Trager's integration algorithm, we need to define some concepts from the theory of algebraic curves. Given a finite algebraic extension $E = K(z)(y)$ of $K(z)$, a *place* P of E is a proper local subring of E containing K, and a *divisor* is a formal sum $\sum n_P P$ with finite support, where the n_P's are integers and the

P's are places. Let P be a place, then its maximal ideal μ_P is principal, so let $p \in E$ be a generator of μ_P. The *order at P* is the function $\nu_P : E^* \to \mathbb{Z}$ which maps $f \in E^*$ to the largest $k \in \mathbb{Z}$ such that $f \in p^k P$. Given $f \in E^*$, the *divisor of f* is $(f) = \sum \nu_P(f) P$ where the sum is taken over all the places. It has finite support since $\nu_P(f) \neq 0$ if and only if P is a pole or zero of f. Finally, we way that a divisor $\delta = \sum n_P P$ is *principal* if $\delta = (f)$ for some $f \in E^*$. Note that if δ is principal, the $\sum n_P = 0$, but the converse is not generally true, except if $E = K(z)$. Trager's algorithm proceeds essentially by constructing candidate divisors for the u_i's of (14):

- Let $\sum_{i=1}^n A_i w_i$ be the numerator of h with respect to w, and D be its (squarefree) denominator
- Write $\sum_{i=1}^n A_i w_i = G/H$, where $G \in K[z, y]$ and $H \in K[z]$
- Let $f \in K[z, y]$ be the (monic) minimum polynomial for y over $K(z)$, t be a new indeterminante and compute
$$R(t) = \mathrm{resultant_z}\left(\mathrm{pp_t}\left(\mathrm{resultant_y}\left(G - tH\frac{dD}{dz}, F\right)\right), D\right) \in K[t]$$
- Let $\alpha_1, \ldots, \alpha_s \in \overline{K}$ be the distinct nonzero roots of R, $(q_1, \ldots, q_k)$ be a basis for the vector space that they generate over $\mathbb{Q}$, write $\alpha_i = r_{i1} q_1 + \cdots + r_{ik} q_k$ for each i, where $r_{ij} \in \mathbb{Q}$ and let $m > 0$ be a common denominator for all the r_{ij}'s
- For $1 \le j \le k$, let $\delta_j = \sum_{i=1}^s m r_{ij} \sum_l r_l P_l$ where r_l is the ramification index of P_l and P_l runs over all the places at which $h\,dz$ has residue $r_i \alpha_i$
- If there are nonzero integers $n_1, \ldots, n_k$ such that $n_j \delta_j$ is principal for each j, then let
$$u = h - \frac{1}{m} \sum_{j=1}^k \frac{q_j}{n_j u_j} \frac{du_j}{dz}$$
where $u_j \in E(\alpha_1, \ldots, \alpha_s)^*$ is such that $n_j \delta_j = (u_j)$. If $u = 0$, then $\int h\,dz = \sum_{j=1}^k q_j \log(u_j)/(m n_j)$, otherwise if either $u \neq 0$ or there is no such integer n_j for at least one j, then $h\,dz$ has no elementary integral.

Note that this algorithm expresses the integral, when it is elementary, with the smallest possible number of logarithms. Steps 3 to 6 requires computing in the splitting field K_0 of R over K, but it can be proven that, as in the case of rational functions, K_0 is the minimal algebraic extension of K necessary to express the integral in the form (4). Trager [Tr84] describes a representation of divisors as fractional ideals and gives algorithms for the arithmetic of divisors and for testing whether a given divisor is principal. In order to determine whether there exists an integer N such that $N\delta$ is principal, we need to reduce the algebraic extension to one over a finite field $\mathbb{F}_{p^q}$ for some "good" prime $p \in \mathbb{Z}$. Over

$\mathbb{F}_{p^q}$, it is known that for every divisor $\delta = \sum n_P P$ such that $\sum n_P = 0$, $M\delta$ is principal for some integer $1 \le M \le (1+\sqrt{p^q})^{2g}$, where g is the genus of the curve [We71], so we compute such an M by testing $M = 1, 2, 3, \ldots$ until we find it. It can then be shown that for almost all primes p, if $M\delta$ is not principal in characteristic 0, the $N\delta$ is not principal for any integer $N \neq 0$. Since we can test whether the prime p is "good" by testing whether the image in $\mathbb{F}_{p^q}$ of the discriminant of the discriminant of the minimal polynomial for y over $K[z]$ is 0, this yields a complete algorithm. In the special case of hyperelliptic extensions, *i.e.* simple radical extensions of degree 2, Bertrand [Ber95] describes a simpler representation of divisors for which the arithmetic and principality tests are more efficient than the general methods.

Example 4 *Continuing example 3, we were left with the integrand*

$$\frac{\sqrt{x^8+1}}{x(x^8+1)} = \frac{w_2}{x(x^8+1)} \in E = \mathbb{Q}(x)[y]/(y^2 - x^8 - 1)$$

where $(w_1, w_2) = (1, y)$ is an integral basis normal at infinity, and the denominator $D = x(x^8+1)$ of the integrand is squarefree. Its numerator is $w_2 = y$, so the resultant of step 3 is

$$\mathit{resultant}_x(\mathit{pp}_t(\mathit{resultant}_y(y - t(9x^8+1), y^2 - x^8 - 1)), x(x^8+1)) = ct^{16}(t^2-1)$$

where c is a large nonzero integer. Its nonzero roots are ± 1, and the integrand has residue 1 at the place P corresponding to the point $(x, y) = (0, 1)$ and -1 at the place Q corresponding to the point $(x, y) = (0, -1)$, so the divisor δ_1 of step 5 is $\delta_1 = P - Q$. It turns out that δ_1, $2\delta_1$, and $3\delta_1$ are not principal, but that

$$4\delta_1 = \left(\frac{x^4}{1+y}\right) \quad \text{and} \quad \frac{w_2}{x(x^8+1)} - \frac{1}{4}\frac{(x^4/(1+y))'}{x^4/(1+y)} = 0$$

which implies that

$$\int \frac{\sqrt{x^8+1}}{x(x^8+1)}\, dx = \frac{1}{4}\log\left(\frac{x^4}{1+\sqrt{x^8+1}}\right)$$

Example 5 *Consider*

$$\int \frac{x\, dx}{\sqrt{1-x^3}}$$

The integrand is

$$f = \frac{xy}{1-x^3} \in E = \mathbb{Q}(x)[y]/(y^2 + x^3 - 1)$$

where $(w_1, w_2) = (1, y)$ is an integral basis normal at infinity, and the denominaotr $D = 1 - x^3$ of the integrand is squarefree. Its numerator is $xw_2 = xy$, so the resultant of step 3 is

$$\mathit{resultant}_x(\mathit{pp}_t(\mathit{resultant}_y(xy + 3tx^2, y^2 + x^3 - 1)), 1 - x^3) = 729t^6$$

whose only root is 0. Since $f \neq 0$, *we conclude from step 6 that* $\int f\,dx$ *is not an elementary function.*

Example 6

$$\int \frac{dx}{x\sqrt{1-x^3}}$$

The integrand is

$$f = \frac{y}{x - x^4} \in E = \mathbb{Q}(x)[y]/(y^2 + x^3 - 1)$$

where $(w_1, w_2) = (1, y)$ *is an integral basis normal at infinity, and the denominator* $D = x - x^4$ *of the integrand is squarefree. Its numerator is* $w_2 = y$, *so the resultant of step 3 is*

$$resultant_x(pp_t(resultant_y(y + t(4x^3 - 1), y^2 + x^3 - 1)), x - x^4) = 729t^6(t^2 - 1)$$

Its nonzero roots are ± 1, *and the integrand has residue 1 at the place* P *corrsesponding to the point* $(x, y) = (0, 1)$ *and* -1 *at the place* Q *corresponding to the point* $(x, y) = (0, -1)$ *so the divisor* δ_1 *of step 5 is* $\delta_1 = P - Q$. *It turns out that* δ_1 *and* $2\delta_1$ *are not principal, but that*

$$3\delta_1 = \left(\frac{y-1}{y+1}\right) \quad \text{and} \quad \frac{y}{x - x^4} - \frac{1}{3}\frac{((y-1)/(y+1))'}{(y-1)/(y+1)} = 0$$

which implies that

$$\int \frac{dx}{x\sqrt{1-x^3}} = \frac{1}{3}\log\left(\frac{\sqrt{1-x^3}-1}{\sqrt{1-x^3}+1}\right)$$

1.3 Elementary Functions

Let f be an arbitrary elementary function. In order to generalize the algorithms of the previous sections, we need to build an algebraic model in which f behaves in some sense like a rational or algebraic function. For that purpose, we need to formally define differential fields and elementary functions.

1.3.1 Differential algebra

A *differential field* $(K,')$ is a *differential extension* of $(K,')$ with a given map $a \to a'$ from K into K, satisfying $(a + b)' = a' + b'$ and $(ab)' = a'b + ab'$. Such a map is called a *derivation* on K. An element $a \in K$ which satisfies $a' = 0$ is called a *constant*, and the set Const(K)$= \{a \in K$ such that $a' = 0\}$ of all the constants of K is a subfield of K.

A differential field $(E,')$ is a *differential equation* of $(K,')$ if $K \subseteq E$ and the derivation on E extends the one on K. In that case, an element $t \in E$ is a *monomial* over K if t is transcendental over K and $t' \in K[t]$, which implies that both $K[t]$ and $K(t)$ are closed under '. An element $t \in E$ is *elementary over* K if either

- $t' = b'/b$ for some $b \in K^*$, in which case we say that t is a *logarithm* over K, and write $t = log(b)$, or
- $t' = b't$ for some $b \in K^*$, in which case we say that t is an *exponential* over K, and write $t = e^b$, or
- t is algebraic over K

A differential extension $(E,')$ of $(K,')$ is *elementary over* K, if there exist $t_1, \dots, t_m$ in E such that $E = K(t_1, \dots, t_m)$ and each t_i is elementary over $K(t_1, \dots, t_{i-1})$. We say that $f \in K$ *has an elementary integral* over K if there exists an elementary extension $(F,')$ of $(K,')$ and $g \in F$ such that $g' = f$. An *elementary function* of the variable x is an element of an elementary extension of the rational function field $(C(x), d/dx)$, where $C = \text{Const}(C(x))$.

Elementary extensions are useful for modeling any function as a rational or algebraic function of one main variable over the other terms present in the function: given an elementary integrand $f(x)\ dx$, the integration algorithm first constructs a field C containing all the constants appearing in f, then the rational function field $(C(x), d/dx)$, then an elementary tower $E = C(x)(t_1, \dots, t_k)$ containing f. Note that such a tower is not unique, and in addition, ajoining a logarithm could in fact adjoin a new constant, and an exponential could in fact be algebraic, for example $\mathbb{Q}(x)(log(x), log(2x)) = \mathbb{Q}(log(2))(x)(log(x))$ and $\mathbb{Q}(x)(e^{log(x)/2}) = \mathbb{Q}(x)(\sqrt{x})$. There are however algorithms that detect all such occurences and modify the tower accordingly [Ris79], so we can assume that all the logarithms and exponentials appearing in E are monomials, and that $\text{Const}(E) = C$. Let now k_0 be the largest index such that t_{k_0} is transcendental over $K = C(x)(t_1, \dots, t_{k_0-1})$ and $t = t_{k_0}$. Then E is a finitely generated algebraic extension of $K(t)$, and in the special case $k_0 = k$, $E = K(t)$. Thus, $f \in E$ can be seen as a univariate rational or algebraic function over K, the major difference with the pure rational or algebraic cases being that K is not constant with respect to the derivation. It turns out that the algorithms of the previous section can be generalized to such towers, new methods being required only for the polynomial (or integral) part. We note that Liouville's Theorem remains valid when E is an arbitrary differential field, so the integration algorithms work by attempting to solve equation (13) as previously.

Example 7 *The function (1) is the element* $f = (t - t^{-1})\sqrt{-1}/2$ *of* $E = K(t)$ *where* $K = \mathbb{Q}(\sqrt{-1})(x)(t_1, t_2)$ *with*

$$t_1 = \sqrt{x^3 - x + 1}, \quad t_2 = e^{2\sqrt{-1}(x^3 - t_1)}, \quad \text{and} \quad t = e^{((1-t_2)/(1+t_2)) - x\sqrt{-1}}$$

which is transcendental over K. Alternatively, it can also be written as the element $f = 2\theta/(1+\theta^2)$ of $F = K(\theta)$ where $K == \mathbb{Q}(x)(\theta_1, \theta_2)$ with

$$\theta_1 = \sqrt{x^3 - x + 1}, \quad \theta_2 = \tan(x^3 - \theta_1), \quad \text{and} \quad \theta = \tan\left(\frac{x+\theta_2}{2}\right)$$

which is a transcendental monomial over K. It turns out that both towers can be used in order to integrate f.

The algorithms of the previous sections relied extensively on squarefree factorization and on the concept of squarefree polynomials. The appropriate analogue in monomial extensions is the notion of *normal* polynomials: let t be a monomial over K, we say that $p \in K[t]$ is *normal* (with respect to ') if $\gcd(p, p') = 1$, and that p is *special* if $\gcd(p, p') = p$, *i.e.* $p|p'$ in $K[t]$. For $p \in K[t]$ squarefree, let $p_s = \gcd(p, p')$ and $p_n = p/p_s$. Then $p = p_s p_n$, while p_s is special and p_n is normal. Therefore, squarefree factorization can be used to write any $q \in K[t]$ as a product $q = q_s q_n$, where $\gcd(q_s, q_n) = 1$, q_s is special and all the squarefree factors of q_n are normal. We call q_s the *special part* of q and q_n its *normal part*.

1.3.2 The Hermite reduction

The Hermite reductions we presented for rational and algebraic functions work in exactly the same way algebraic extensions of monomial extensions of K, as long as we apply them only to the normal part of the denominator of the integrand. Thus, if D is the denominator of the integrand, we let S be the special part of D, $D_1 D_2^2 \dots D_m^m$ be a squarefree factorization of the *normal* part of D, $V = D_m$, $U = D/V^m$ and the rational and algebraic Hermite reductions proceed normally, eventually yielding an integrand whose denominator has a squarefree normal part.

Example 8 *Consider*

$$\int \frac{x - \tan(x)}{\tan(x)^2}\, dx$$

The integrand is

$$f = \frac{x-t}{t^2} \in K(t) \quad \text{where } K = \mathbb{Q}(x) \text{ and } t' = t^2 + 1$$

Its denominator is $D = t^2$, and $\gcd(t, t') = 1$ implying that t is normal, so $m = 2$, $V = t$, $U = D/t^2 = 1$, and the extended Euclidean algorithm yields

$$\frac{A}{1-m} = t - x = -x(t^2+1) + (xt+1)t = -xUV' + (xt+1)V$$

implying that

$$\int \frac{x - \tan(x)}{\tan(x)^2}\, dx = -\frac{x}{\tan(x)} - \int x\, dx$$

and the remaining integrand has a squarefree denominator.

Example 9 *Consider*

$$\int \frac{\log(x)^2 + 2x\log(x) + x^2 + (x+1)\sqrt{x+\log(x)}}{x\log(x)^2 + 2x^2\log(x) + x^3}\,dx$$

The integrand is

$$f = \frac{t^2 + 2xt + x^2 + (x+1)y}{xt^2 + 2x^2t + x^3} \in E = K(t)[y]/(y^2 - x - t)$$

where $K = \mathbb{Q}(x)$ and $t = log(x)$. The denominator of f with respect to the basis $w = (1, y)$ is $D = xt^2 + 2x^2t + x^3$ whose squarefree factorization is $x(t+x)^2$. Both x and $t + x$ are normal, so $m = 2$, $V = t + x$, $U = D/V^2 = x$, and the solution (12) of (8) is

$$f_1 = \frac{t^2 + 2xt + x^2}{x(-(t'+1))} = -\frac{t^2 + 2xt + x^2}{x+1},$$

$$f_2 = \frac{x+1}{x\left((t+x)\frac{1}{2}\frac{t'+1}{t+z} - (t'+1)\right)} = -2$$

We have $Q = 1$, so $0V + 1Q = 1$, $A = 0$, $R = 1$, $RQf_1 = f_1 = -V^2/(x+1)$ and $RQf_2 = f_2 = 0V - 2$, so $B = -2y$ and

$$h = f - \left(\frac{B}{V}\right)' = \frac{1}{x}$$

implying that

$$\int \frac{\log(x)^2 + 2x\log(x) + x^2 + (x+1)\sqrt{x+\log(x)}}{x\log(x)^2 + 2x^2\log(x) + x^2}\,dx = \frac{2}{\sqrt{x + log(x)}} + \int \frac{dx}{x}$$

and the remaining integrand has a squarefree denominator.

1.3.3 The polynomial reduction

In the transcendental case $E = K(t)$ and when t is a monomial satisfying $deg_t(t') \geq 2$, then it is possible to reduce the degree of the polynomial part of the integrand until it is smaller than $deg_t(t')$. In the case when $t = \tan(b)$ for some $b \in K$, then it is possible either to prove that the integral is not elementary, or to reduce the polynomial part of the integrand to be in K. Let $f \in K(t)$ be our integrand and write $f = P + A/D$, where $P, A, D \in K[t]$ and $deg(A) < deg(D)$. Write $P = \sum_{i=1}^{e} p_i t^i$ and $t' = \sum_{i=0}^{d} c_i t^i$ where $p_0, \ldots, p_e, c_0, \ldots, c_d \in K$, $d \geq 2$, $p_e \neq 0$ and $c_d \neq 0$. It is easy to verify that if $e \geq d$, then

$$P = \left(\frac{a_e}{(e-d+1)c_d} t^{e-d_1}\right)' + \overline{P} \tag{1.15}$$

where $\overline{P} \in K[t]$ is such that $\overline{P} = 0$ or $deg_t(\overline{P}) < e$. Repeating the above transformation we obtain $Q, R \in K[t]$ such that $R = 0$ or $deg_t(R) < d$ and $P = Q' + R$. Write then $R = \sum_{i=0}^{d-1} r_i t^i$ where $r_0, \ldots, r_{d-1} \in K$. Again, it is easy to verify that for any special $S \in K[t]$ with $deg_t(S) > 0$, we have

$$R = \frac{1}{deg_t(S)} \frac{r_{d-1}}{c_d} \frac{S'}{S} + \overline{R}$$

where $\overline{R} \in K[t]$ is such that $\overline{R} = 0$ or $deg_t(\overline{R}) < e - 1$. Furthermore, it can be proven [Bro97] that if $R + A/D$ has an elementary integral over $K(t)$, then r_{d-1}/c_d is a constant, which implies that

$$\int R = \frac{1}{deg_t(S)} \frac{r_{d-1}}{c_d} \log(S) + \int \left(\overline{R} + \frac{A}{D} \right)$$

so we are left with an integrand whose polynomial part has degree at most $deg_t(t') - 2$. In this case $t = \tan(b)$ for $b \in K$, then $t' = b't^2 + b'$, so $\overline{R} \in K$.

Example 10 *Consider*

$$\int (1 + x\tan(x) + \tan(x)^2)\ dx$$

The integrand is

$$f = 1 + xt + t^2 \in K(t) \quad \text{where } K = \mathbb{Q}(x) \text{ and } t' = t^2 + 1$$

Using (15), we get $\overline{P} = f - t' = f - (t^2 + 1) = xt$ *so*

$$\int (1 + x\tan(x) + \tan(x)^2)\ dx = \tan(x) + \int x\tan(x)\ dx$$

and since $x' \neq 0$, *the above criterion imples that the remaining integral is not an elementary function.*

1.3.4 The residue criterion

Similarly to the Hermite reduction, the Rothstein-Trager and Lazard-Rioboo-Trager algorithms are easy to generalize to the transcendental case $E = K(t)$ for arbitrary monomials t: let $f \in K(t)$ be our integrand and write $f = P + A/D + B/S$ where $P, A, D, B, S \in K[t]$, $deg(A) < deg(D)$, S is special and, following the Hermite reduction, D is normal. Let then z be a new indeterminate, $\kappa : K[z] \to K[z]$ be give by $\kappa(\sum_i a_i z^i) = \sum_i a_i' z^i$,

$$R = \mathrm{resultant_t}(D, A - zD') \in K[z]$$

be the Rothstein-Trager resultant, $R = R_1 R_2^2 \ldots R_k^k$ be its squarefree factorization, $Q_i = \gcd_z(R_i, \kappa(R_i))$ for each i, and

$$g = \sum_{i=1}^{k} \sum_{a | Q_i(a) = 0} a \log(\gcd{}_t(D, A - aD'))$$

Note that the roots of each Q_i must all be constants, and that the arguments of the logarithms can be obtained directly from the subresultant PRS of D and $A - zD'$ as in the rational function case. It can then be proven [Bro97] that

- $f - g'$ is always "simpler" than f
- the splitting field of $Q_1 \cdots Q_k$ over K is the minimal algebraic extension of K needed in order to express $\int f$ in the form (4)
- if f has an elementary integral over $K(t)$, then $R|\kappa(R)$ in $K[z]$ and the denominator of $f - g'$ is special

Thus, while in the pure rational function case the remaining integrand is a polynomial, in this case the remaining integrand has a special denominator. In that case we have additionally that if its integral is elementary, then (13) has a solution such that $v \in K(t)$ has a special denominator, and each $u_i \in K(c_1, \ldots, c_k)[t]$ is special.

Example 11 *Consider*

$$\int \frac{2\log(x)^2 - \log(x) - x^2}{\log(x)^3 - x^2\log(x)}\, dx$$

The integrand is

$$f = \frac{2t^2 - t - x^2}{t^2 - xt^2} \in K(t) \quad \text{where } K = \mathbb{Q}(x) \text{ and } t = \log(x)$$

Its denominator is $D = t^3 - x^2 t$, which is normal, and the resultant is

$$\begin{aligned} R &= \mathit{resultant}_t\left(t^3 - x^2 t, \frac{2x - 3z}{x} t^2 + (2xz - 1)t + x(z - x)\right) \\ &= 4x^3(1 - x^2)\left(z^3 - xz^2 - \frac{1}{4}z + \frac{x}{4}\right) \end{aligned}$$

which is squarefree in $K[z]$. We have

$$\kappa(R) = -x^2(4(5x^2 + 3)z^3 + 8x(3x^2 - 2)z^2 + (5x^2 - 3)z - 2x(3x^2 - 2))$$

so

$$Q_1 = \gcd{}_z(R, \kappa R) = x^2\left(z^2 - \frac{1}{4}\right)$$

and

$$\gcd{}_t\left(t^3 + x^2 t, \frac{2x - 3a}{x}t^2 + (2xa - 1)t + x(a - x)\right) = t + 2ax$$

where $a^2 - 1/4 = 0$, whence

$$g = \sum_{a|a^2 - 1/4 = 0} a\log(t + 2ax) = \frac{1}{2}\log(t + x) - \frac{1}{2}\log(t - x)$$

Computing $f - g'$ we find

$$\int \frac{2\log(x)^2 - \log(x) - x^2}{\log(x)^3 - x^2\log(x)}\,dx = \frac{1}{2}\log\left(\frac{\log(x)+x}{\log(x)-x}\right) + \int \frac{dx}{\log(x)}$$

and since $deg_z(Q_1) < deg_z(R)$, it follows that the remaining integral is not an elementary function (it is in fact the logarithmic integral $Li(x)$).

In the most general case, when $E = K(t)(j)$ is algebraic over $K(t)$ and y is integral over $K[t]$, the criterion part of the above result remains valid: let $w = (w_1, \ldots, w_n)$ be an integral basis for E over $K(t)$ and write the integrand $f \in E$ as $f = \sum_{i=1}^n A_i w_i/D + \sum_{i=1}^n B_i w_i/S$ where S is special and, following the Hermite reduction, D is normal. Write $\sum_{i=1}^n A_i w_i = G/H$, where $G \in K[t,y]$ and $H \in K[t]$, let $F \in K[t,y]$ be the (monic) minimum polynomial for y over $K(t)$, z be a new indeterminante and compute

$$R(z) = \mathrm{resultant_t}(\mathrm{pp_z}(\mathrm{resultant_y}(G - tHD', F)), D) \in K[t] \qquad (1.16)$$

It can then be proven [Bro90] that if f has an elementary integral over E, then $R|\kappa(R)$ in $K[z]$.

Example 12 *Consider*

$$\int \frac{\log(1+e^x)^{(1/3)}}{1+\log(1+e^x)}\,dx \qquad (1.17)$$

The integrand is

$$f = \frac{y}{t+1} \in E = K(t)[y]/(y^3 - t)$$

where $K = \mathbb{Q}(x)(t_1)$,$t_1 = e^x$ and $t = \log(1+t_1)$. Its denominator with respect to the integral basis $w = (1, y, y^2)$ is $D = t+1$, which is normal, and the resultant is

$$R = \mathrm{resultant_t}(\mathrm{pp_z}(\mathrm{resultant_y}(y - zt_1/(1+t_1), y^3 - t)), t+1) = -\frac{t_1^3}{(1+t_1)^3}z^3 - 1$$

We have

$$\kappa(R) = -\frac{3t_1^3}{(1+t_1)^4}z^3$$

which is coprime with R in $K[z]$, implying that the integral (17) is not an elementary function.

1.3.5 The transcendental logarithmic case

Suppose now that $t = \log(b)$ for some $b \in K^*$, and that $E = K(t)$. Then, every special polynomial must be in K, so, following the residue criterion, we must look for a solution $v \in K[t]$, $u_1, \ldots, u_k \in K(c_1, \ldots, c_n)^*$ of (13). Furthermore, the integrand f is also in $K[t]$, so write $f = \sum_{i=0}^{d} f_i t^i$ where $f_0, \ldots, f_d \in K$ and $f_d \neq 0$. We must have $deg_t(v) \leq d_1$, so writing $v = \sum_{i=0}^{d+1} v_i t^i$, we get

$$\int f_d t^d + \cdots + f_1 t + f_0 = v_{d+1} t^{d+1} + \cdots + v_1 t + v_0 + \sum_{i=1}^{k} c_i \log(u_i)$$

If $d = 0$, then the above is simply an integration problem for $f_0 \in K$, which can be solved recursively. Otherwise, differentiating both sides and equating the coefficients of t^d, we get ${v_{d+1}}' = 0$ and

$$f_d = v_d' + (d+1) v_{d+1} \frac{b'}{b} \tag{1.18}$$

Since $f_d \in K$, we can recursively apply the integration algorithm to f_d, either proving that (18) has no solution, in which case f has no elementary integral, or obtaining the constant v_{d+1}, and v_d up to an additive constant (in fact, we apply recursively a specialized version of the integration algorithm to equations of the form (18), see [Bro97] for details). Write then $v_d = \overline{v_d} + c_d$ where $\overline{v_d} \in K$ is known and $c_d \in \mathrm{Const}(K)$ is undetermined. Equating the coefficients of t^{d-1} yields

$$f_{d-1} - d\overline{v_d}\frac{b'}{b} = {v_{d-1}}' + d c_d \frac{b'}{b}$$

which is an equation of the form (18), so we again recursively compute c_d and v_{d-1} up to an additive constant. We repeat this process until either one of the recursive integrations fails, in which case f has no elementary integral, or we reduce our integrand to an element of K, which is then integrated recursively. The algorithm of this section can also be applied to real arc-tangent extensions, *i.e.* $K(t)$ where t is a monomial satisfying $t' = b'/(1+b^2)$ for some $b \in K$.

1.3.6 The transcendental exponential case

Suppose now that $t = e^b$ for some $b \in K$, and that $E = K(t)$. Then, every nonzero special polynomial must be of the form at^m for $a \in K^*$ and $m \in \mathbb{N}$. Since

$$\frac{(at^m)'}{at^m} = \frac{a'}{a} + m\frac{t'}{t} = \frac{a'}{a} + mb'$$

we must then look for a solution $v \in K[t, t^{-1}]$, $u_1, \ldots, u_k \in K(c_1, \ldots, c_n)^*$ of (13). Furthermore, the integrand f is also in $K[t, t^{-1}]$, so write $f = \sum_{i=e}^{d} f_i t^i$ where $f_e, \ldots, f_d \in K$ and $e, d \in \mathbb{Z}$. Since $(at^m)' = (a' + mb')t^m$ for any $m \in \mathbb{Z}$, we must have $v = Mb + \sum_{i=e}^{d} v_i t^i$ for some integer M, hence

$$\int \sum_{i=e}^{d} f_i t^i = Mb + \sum_{i=e}^{d} v_i t^i + \sum_{i=1}^{k} c_i \log(u_i)$$

Differentiating both sides and equating the coefficients of each power to t^d, we get

$$f_0 = (v_0 + Mb)' + \sum_{i=1}^{k} c_i \frac{u_i'}{u_i}$$

which is simply an integration problem for $f_0 \in K$, and

$$f_i = v_i' + ib'v_i \quad \text{for } e \le i \le d, i \ne 0$$

The above problem is called a *Risch differential equation overK*. Although solving it seems more complicated than solving $g' = f$, it is actually simpler than an integration problem because we look for the solutions v_i in K only rather than in an extension of K. Bronstein [Bro90, Bro91, Bro97] and Risch [Ris68, Ris69a, Ris69b] describe algorithms for solving this type of equation when K is an elementary extension of the rational function field.

1.3.7 The transcendental tangent case

Suppose now that $t = \tan(b)$ for some $b \in K$, *i.e.* $t' = b'(1+t^2)$, that $\sqrt{-1} \notin K$ and that $E = K(t)$. Then, every nonzero special polynomial must be of the form $a(t^2+1)^m$ for $a \in K^*$ and $m \in \mathbb{N}$. Since

$$\frac{(a(t^2+1)^m)'}{a(t^2+1)^m} = \frac{a'}{a} + m\frac{(t^2+1)'}{t^2+1} = \frac{a'}{a} + 2mb't$$

we must look for $v = V/(t^2+1)^m$ where $V \in K[t]$, $m_1, \ldots, m_k \in \mathbb{N}$, constants $c_1, \ldots, c_k \in \overline{K}$ and $u_1, \ldots, u_k \in K(c_1, \ldots, c_k)^*$ such that

$$f = v' + 2b't\sum_{i=1}^{k} c_i m_i + \sum_{i=1}^{k} c_i \frac{u_i'}{u_i}$$

Furthermore, the integrand $f \in K(t)$ following the residue criterion must be of the form $f = A/(t^2+1)^M$ where $A \in K[t]$ and $M \ge 0$. If $M > 0$, it can be shown that $m = M$ and that

$$\begin{pmatrix} c' \\ d' \end{pmatrix} + \begin{pmatrix} 0 & -2mb' \\ 2mb' & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \tag{1.19}$$

where $at+b$ and $ct+d$ are the remainders module t^2+1 of A and V respectively. The above is a coupled differential system, which can be solved by methods similar to the ones used for Risch differential equations [Bro97]. If it has no solution, then the integral is not elementary, otherwise we reduce the integrand to $h \in K[t]$, at which point the polynomial reduction either proves that its integral is not elementary, or reduce the integrand to an element of K, which is integrated recursively.

Example 13 *Consider*

$$\int \frac{sin(x)}{x}\, dx$$

The integrand is

$$f = \frac{2t/x}{t^2+1} \in K(t) \quad \text{where } K = \mathbb{Q}(x) \text{ and } t = \tan\left(\frac{x}{2}\right)$$

Its denominator is $D = t^2 + 1$, which is special, and the system (19) becomes

$$\begin{pmatrix} c' \\ d' \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 2/x \\ 0 \end{pmatrix}$$

which has no solution in $\mathbb{Q}(x)$, implying that the integral is not an elementary function.

1.3.8 The algebraic logarithmic case

The transcendental logarithmic case method also generalizes to the case when $E = K(t)(y)$ is algebraic over $K(t)$, $t = log(b)$ for $b \in K^*$ and y is integral over $K[t]$: following the residue criterion, we can assume that $R|\kappa(R)$ where R is given by (16), hence that all its roots in $\overline{K}$ are constants. The polynomial part of the integrand is replace by a family of at most $[E : K(t)]$ Puiseux expansions at infinity, each of the form

$$a_{-m}\theta^{-m} + \cdots + a_{-1}\theta^{-1} + \sum_{i \geq 0} a_i \theta^i \tag{1.20}$$

where $\theta^r = t^{-1}$ for some positive integer r. Applying the integration algorithm recursively to $a_r \in \overline{K}$, we can test whether there exist $\rho \in \text{Const}(\overline{K})$ and $v \in \overline{K}$ such that

$$a_r = v' + \rho \frac{b'}{b}$$

If there are no such v and c for at least one of the series, then the integral is not elementary, otherwise ρ is uniquely determined by a_r, so let $\rho_1, \ldots, \rho_q$ where $q \leq [E : K(t)]$ be the distinct constants we obtain, $\alpha_1, \ldots, \alpha_s \in \overline{K}$ be the distinct nonzero roots of R, and $(q_1, \ldots, q_k)$ be a basis for the vector space generated by the ρ_i's and α_i's over $\mathbb{Q}$. Write $\alpha_i = r_{i1}q_1 + \cdots + r_{ik}q_k$ and $\rho_i = s_{i1}q_1 + \cdots + s_{ik}q_k$ for each i, where $r_{ij}, s_{ij} \in \mathbb{Q}$ and let $m > 0$ be a common denominator for all the r_{ij}'s and s_{ij}'s. For $1 \leq j \leq k$, let

$$\delta_j = \sum_{i=1}^{s} m r_{ij} \sum_l r_l P_l - \sum_{i=1}^{q} m s_{ij} \sum_l s_l Q_l$$

where r_l is the ramification index of P_l, s_l is the ramification index of Q_l, P_l runs over all the finite places at which $h\ dz$ has residue $r_l\alpha_i$ and Q_l runs over all the infinite places at which $\rho = \rho_i$. As in the pure algebraic case, if there is a j for which $N\delta_j$ is not principal for any nonzero integer N, then the integral is not elementary, otherwise, let $n_1, \dots, n_k$ be nonzero integers such that $n_j\delta_j$ is principal for each j, and

$$h = f - \frac{1}{m}\sum_{j=1}^{k}\frac{q_j}{n_j}\frac{u_j^{'}}{u_j}$$

where f is the integrand and $u_j \in E(\alpha_1, \dots, \alpha_s, \rho_1, \dots, \rho_q)^*$ is such that $n_j\delta_j = (u_j)$. If the integral of h is elementary, then (13) must have a solution with $v \in \mathbf{O}_{K[x]}$ and $u_1, \dots, u_k \in \overline{K}$ so we must solve

$$h = \frac{\sum_{i=1}^{n} A_i w_i}{D} = \sum_{i=1}^{n} v_i^{'} w_i + \sum_{i=1}^{n} v_i w_i^{'} + \sum_{i=1}^{k} c_i \frac{u_i^{'}}{u_i} \tag{1.21}$$

for $v_1, \dots, v_n \in K[t]$, constants $c_1, \dots, c_n \in \overline{K}$ and $u_1, \dots, u_k \in \overline{K}^*$ where $w = (w_1, \dots, w_n)$ is an integral basis for E over $K(t)$.

If E is a simple radical extension of $K(t)$, and we use the basis (11) and the notation of that section, then $w_1 = 1$ and

$$w_i^{'} = \left(\frac{i-1}{n}\frac{H'}{H} - \frac{D_{i-1}^{'}}{D_{i-1}}\right) w_i \quad \text{for } 1 \le i \le n \tag{1.22}$$

This implies that (21) becomes

$$\frac{A_1}{D} = v_1^{'} + \sum_{i=1}^{k} c_i \frac{u_i^{'}}{u_i} \tag{1.23}$$

which is simply an integration problem for $A_1/D \in K(t)$, and

$$\frac{A_i}{D} = v_i^{'} + \left(\frac{i-1}{n}\frac{H'}{H} - \frac{D_{i-1}^{'}}{D_{i-1}}\right) v_i \quad \text{for } 1 < i \le n \tag{1.24}$$

which are Risch differential equations over $K(t)$

Example 14 *Consider*

$$\int \frac{(x^2+2x+1)\sqrt{x+\log(x)} + (3x+1)\log(x) + 3x^2 + x}{(x\log(x)+x^2)\sqrt{x+\log(x)} + x^2\log(x) + x^3}\, dx$$

The integrand is

$$f = \frac{((3x+1)t - x^3 + x^2)y - (2x^2 - x - 1)t - 2x^3 + x^2 + x}{xt^2 - (x^3 - 2x^2)t - x^4 + x^3} \in E = K(t)[y]/(F)$$

where $F = y^2 - x - t$, $K = \mathbb{Q}(x)$ *and* $t = \log(x)$. *Its denominator with respect to the integral basis* $w = (1, y)$ *is* $D = xt^2 - (x^3 - 2x^2)t - x^4 + x^3$, *which is normal, and the resultant is*

$$\begin{aligned} R &= \mathrm{resultant_t}(\mathrm{pp_z}(\mathrm{resultant_y}(((3x+1)t - x^3 + x^2)y \\ &\qquad - (2x^2 - x - 1)t - 2x^3 + x^2 + x - zD', F)), D) \\ &= x^{12}(2x+1)^2(x+1)^2(x-1)^2 z^3(z-2) \end{aligned}$$

We have

$$\kappa(R) = \frac{36x^3 + 16x^2 - 28x - 12}{x(2x+1)(x+1)(x-1)} R$$

so $R|\kappa(R)$ *in* $K[z]$. *Its only nonzero root is 2, and the integrand has residue 2 at the place* P *corresponding to the point* $(t, y) = (x^2 - x, -x)$. *There is only one place* Q *at infinity of ramification index 2, and the coefficient of* t^{-1} *in the Puiseux expansion of* f *at* Q *is*

$$a_2 = 1 - 2x + \frac{1}{x} = (x - x^2)' + \frac{x'}{x}$$

which implies that the corresponding ρ *is 1. Therefore, the divisor for the logand is* $\delta = 2P - 2Q$. *It turns out that* $\delta = (u)$ *where* $u = (x+y)^2 \in E^*$, *so the new integrand is*

$$h = f - \frac{u'}{u} = f - 2\frac{(x+y)'}{x+y} = \frac{(x+1)y}{xt + x^2}$$

We have $y^2 = t + x$, *which is squarefree, so (23) becomes*

$$0 = v_1' + \sum_{i=1}^{k} c_i \frac{u_i'}{u_i}$$

whose solution is $v_1 = k = 0$ *and (24) becomes*

$$\frac{x+1}{xt + x^2} = v_2' + \frac{x+1}{2xt + 2x^2} v_2$$

whose solution is $v_2 = 2$, *implying that* $h = 2y'$, *hence that*

$$\int \frac{(x^2 + 2x + 1)\sqrt{x + \log(x)} + (3x+1)\log(x) + 3x^2 + x}{(x\log(x) + x^2)\sqrt{x + \log(x)} + x^2\log(x) + x^3}\, dx$$

$$2\sqrt{x + \log(x)} + 2\log\left(x + \sqrt{x + \log(x)}\right)$$

In the general case when E is not a radical extension of $K(t)$, (21) is solved by bounding $deg_t(v_i)$ and comparing the Puiseux expansions at infinity of $\sum_{i=1}^{n} v_i w_i$ with those of the form (20) of h, see [Bro90, Ris68] for details.

1.3.9 The algebraic exponential case

The transcendental exponential case method also generalizes to the case when $E = K(t)(y)$ is algebraic over $K(t)$, $t = e^b$ for $b \in K$ and y is integral over $K[t]$: following the residue criterion, we can assume that $R|\kappa(R)$ where R is given by (16), hence that all its roots in $\overline{K}$ are constants. The denominator of the integrancd must be of the form $D = t^m U$ where $\gcd(U,t) = 1$, U is squarefree and $m \geq 0$.

If $m > 0$, E is a simple radical extension of $K(t)$, and we use the basis (11), then it is possible to reduce the power of t appearing in D by a process similar to the Hermite reduction: writing the integrand $f = \sum_{i=1}^n A_i w_i/(t^m U)$, we ask whether we can compute $b_1, \ldots, b_n \in K$ and $C_1, \ldots, C_n \in K[t]$ such that

$$\int \frac{\sum_{i=1}^n A_i w_i}{t^m U} = \frac{\sum_{i=1}^n b_i w_i}{t^m} + \int \frac{\sum_{i=1}^n C_i w_i}{t^{m-1} U}$$

Differentiating both sides and multiplying through by t^m we get

$$\frac{\sum_{i=1}^n A_i w_i}{U} = \sum_{i=1}^n b_i' w_i + \sum_{i=1}^n b_i w_i' - mb' \sum_{i=1}^n b_i w_i + \frac{t \sum_{i=1}^n C_i w_i}{U}$$

Using (22) and equating the coefficients of w_i on both sides, we get

$$\frac{A_i}{U} = b_i' + (\omega_i - mb')b_i + \frac{tC_i}{U} \quad \text{for } 1 \leq i \leq n \tag{1.25}$$

where

$$\omega_i = \frac{i-1}{n}\frac{H'}{H} - \frac{D_{i-1}'}{D_{i-1}} \in K(t)$$

Since $t'/t = b' \in K$, it follows that the denominator of ω_i is not divisible by t in $K[t]$, hence, evaluating (25) at $t = 0$, we get

$$\frac{A_i(0)}{U(0)} = b_i' + (\omega_i(0) - mb')b_i \quad \text{for } 1 \leq i \leq n \tag{1.26}$$

which are Risch differential equations over $K(t)$. If any of them has no solution in $K(t)$, then the integral is not elementary, otherwise we repeat this process until the denominator of the integrand is normal. We then perform the change of variabl $\bar{t} = t^{-1}$, which is also exponential over K since $\bar{t}' = -b'\bar{t}$, and repeat the above process in order to eliminate the power of $\bar{t}$ from the denominator of the integrand. It can be shown that after this process, any solution of (13) must have $v \in K$.

Example 15 *Consider*

$$\int \frac{3(x+e^x)^{(1/3)} + (2x^2+3x)e^x + 5x^2}{x(x+e^x)^{(1/3)}}\, dx$$

The integrand is

$$f = \frac{((2x^2+3x)t+5x^2)y^2+3t+3x}{xt+x^2} \in E = K(t)[y]/(y^3-t-x)$$

where $K = \mathbb{Q}(x)$ *and* $t = e^x$. *Its denominator with respect to the integral basis* $w = (1, y, y^2)$ *is* $D = xt + x^2$, *which is normal, and the resultant is*

$$R = \mathrm{resultant_t}(\mathrm{pp_z}(\mathrm{resultant_y}(((2x^2+3x)t+5x^2)y^2+3t+3x-zD', \\ y^3-t-x)), D) = x^8(1-x)^3z^3$$

We have

$$\kappa(R) = \frac{11x-8}{x(x-1)}R$$

so $R|\kappa(R)$ *in* $K[z]$, *its only root being 0. Since* D *is not divisible by* t, *let* $\overline{t} = t^{-1}$ *and* $z = \overline{t}y$. *We have* $\overline{t}' = -\overline{t}$ *and* $z^3 - \overline{t}^2 - x\overline{t}^3 = 0$, *so the integral basis (11) is*

$$\overline{w} = (\overline{w}_1, \overline{w}_2, \overline{w}_3) = \left(1, z, \frac{z^2}{\overline{t}}\right)$$

Writing f *in terms of that basis gives*

$$f = \frac{3x\overline{t}^2 + 3\overline{t} + (5x^2\overline{t} + 2x^2 + 3x)\overline{w}_3}{x^2\overline{t}^2 + x\overline{t}}$$

whose denominator $\overline{D} = \overline{t}(x + x^2\overline{t})$ *is divisible by* $\overline{t}$. *We have* $H = \overline{t}^2(1 + x\overline{t})$ *so* $D_0 = D_1 = 1$ *and* $D_2 = \overline{t}$, *implying that*

$$\omega_1 = 0, \omega_2 = \frac{(1-3x)\overline{t}-2}{3x\overline{t}+3}, \text{ and } \omega_3 = \frac{(2-3x)\overline{t}-1}{3x\overline{t}+3}$$

Therefore the equations (26) become

$$0 = b_1' + b_1, 0 = b_2' + \frac{1}{3}b_2, \text{ and } 2x+3 = b_3' + \frac{2}{3}b_3$$

whose solutions are $b_1 = b_2 = 0$ *and* $b_3 = 3x$, *implying that the new integrand is*

$$h = f - \left(\frac{3x\overline{w}_3}{\overline{t}}\right)' = \frac{3}{x}$$

hence that

$$\int \frac{3(x+e^x)^{(1/3)} + (2x^2+3x)e^x + 5x^2}{x(x+e^x)^{(1/3)}}\, dx = 3x(x+e^x)^{(2/3)} + 3\int \frac{dx}{x}$$

In the general case when E is not a radical extension of $K(t)$, following the Hermite reduction, any solution of (13) must have $v = \sum_{i=1}^n v_i w_i / t^m$ where $v_1, \ldots, v_m \in K[t]$. We can compute v by bounding $deg_t(v_i)$ and comparing the

Puiseux expansions at $t = 0$ and at infinity of $\sum_{i=1}^{n} v_i w_i / t^m$ with those of the form (20) of the integrand, see [Bro90, Ris68] for details.

Once we are reduced to solving (13) for $v \in K$, constants $c_1, \ldots, c_k \in \overline{K}$ and $u_1, \ldots, u_k \in E(c_1, \ldots, c_k)^*$, constants $\rho_1, \ldots, \rho_s \in \overline{K}$ can be determined at all the places above $t = 0$ and at infinity in a manner similar to the algebraic logarithmic case, at which point the algorithm proceeds by constructing the divisors δ_j and the u_j's as in that case. Again, the details are quite technical and can be found in [Bro90, Ris68, Ris69a].

Chapter 2

Singular Value Decomposition

2.1 Singular Value Decomposition Tutorial

When you browse standard web sources like Wikipedia to learn about Singular Value Decomposition or SVD you find many equations, but not an intuitive explanation of what it is or how it works. SVD is a way of factoring matrices into a series of linear approximations that expose the underlying structure of the matrix. Two important properties are that the linear factoring is exact and optimal. Exact means that the series of linear factors, added together, exactly equal the original matrix. Optimal means that, for the standard means of measuring matrix similarity (the Frobenius norm), these factors give the best possible linear approximation at each step in the series.

SVD is extraordinarily useful and has many applications such as data analysis, signal processing, pattern recognition, image compression, weather prediction, and Latent Sematic Analysis or LSA (also referred to as Latent Semantic Indexing). Why is SVD so useful and how does it work?

As a simple example, let's look at golf scores. Suppose Phil, Tiger, and Vijay play together for 9 holes and they each make par on every hole. Their scorecard, which can also be viewed as a (hole x player) matrix might look like this.

Hole	Par	Phil	Tiger	Vijay
1	4	4	4	4
2	5	5	5	5
3	3	3	3	3
4	4	4	4	4
5	4	4	4	4
6	4	4	4	4
7	4	4	4	4
8	3	3	3	3
9	5	5	5	5

Let's look at the problem of trying to predict what score each player will make on a given hole. One idea is give each hole a HoleDifficulty factor, and each player a PlayerAbility factor. The actual score is predicted by multiplying these two factors together.

PredictedScore = HoleDifficulty * PlayerAbility

For the first attempt, let's make the HoleDifficulty be the par score for the hole, and let's make the player ability equal to 1. So on the first hole, which is par 4, we would expect a player of ability 1 to get a score of 4.

PredictedScore = HoleDifficulty * PlayerAbility = 4 * 1 = 4

For our entire scorecard or matrix, all we have to do is multiply the PlayerAbility (assumed to be 1 for all players) by the HoleDifficulty (ranges from par 3 to par 5) and we can exactly predict all the scores in our example.

In fact, this is the one dimensional (1-D) SVD factorization of the scorecard. We can represent our scorecard or matrix as the product of two vectors, the HoleDifficulty vector and the PlayerAbility vector. To predict any score, simply multiply the appropriate HoleDifficulty factor by the appropriate PlayerAbility factor. Following normal vector multiplication rules, we can

generate the matrix of scores by multiplying the HoleDifficulty vector by the PlayerAbility vector, according to the following equation.

$$\begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ \hline 4 & 4 & 4 \\ 5 & 5 & 5 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \\ 3 & 3 & 3 \\ 5 & 5 & 5 \end{array} = \begin{array}{|c|} 4 \\ 5 \\ 3 \\ 4 \\ 4 \\ 4 \\ 4 \\ 3 \\ 5 \end{array} * \begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ 1 & 1 & 1 \end{array}$$

which is HoleDifficulty * PlayerAbility

Mathematicians like to keep everything orderly, so the convention is that all vectors should be scaled so they have length 1. For example, the PlayerAbility vector is modified so that the sum of the squares of its elements add to 1, instead of the current 12 + 12 + 12 = 3. To do this, we have to divide each element by the square root of 3, so that when we square it, it becomes and the three elements add to 1. Similarly, we have to divide each HoleDifficulty element by the square root of 148. The square root of 3 times the square root of 148 is our scaling factor 21.07. The complete 1-D SVD factorization (to 2 decimal places) is:

$$\begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ \hline 4 & 4 & 4 \\ 5 & 5 & 5 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \\ 3 & 3 & 3 \\ 5 & 5 & 5 \end{array} = \left| \begin{array}{c} 0.33 \\ 0.41 \\ 0.25 \\ 0.33 \\ 0.33 \\ 0.33 \\ 0.33 \\ 0.25 \\ 0.41 \end{array} \right| * \left| \; 21.07 \; \right| * \begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ \hline 0.58 & 0.58 & 0.58 \end{array}$$

which is HoleDifficulty * ScaleFactor * PlayerAbility

Our HoleDifficulty vector, that starts with 0.33, is called the Left Singular Vector. The ScaleFactor is the Singular Value, and our PlayerAbility vector, that starts with 0.58 is the Right Singular Vector. If we represent these 3 parts exactly, and multiply them together, we get the exact original scores. This means our matrix is a rank 1 matrix, another way of saying it has a simple and predictable pattern.

More complicated matrices cannot be completely predicted just by using one set of factors as we have done. In that case, we have to introduce a second set of factors to refine our predictions. To do that, we subtract our predicted scores from the actual scores, getting the residual scores. Then we find a second set of HoleDifficulty2 and PlayerAbility2 numbers that best predict the residual scores.

Rather than guessing HoleDifficulty and PlayerAbility factors and subtracting predicted scores, there exist powerful algorithms than can calculate SVD factorizations for you. Let's look at the actual scores from the first 9 holes of the 2007 Players Championship as played by Phil, Tiger, and Vijay.

Hole	Par	Phil	Tiger	Vijay
1	4	4	4	5
2	5	4	5	5
3	3	3	3	2
4	4	4	5	4
5	4	4	4	4
6	4	3	5	4
7	4	4	4	3
8	3	2	4	4
9	5	5	5	5

The 1-D SVD factorization of the scores is shown below. To make this example easier to understand, I have incorporated the ScaleFactor into the PlayerAbility and HoleDifficulty vectors so we can ignore the ScaleFactor for this example.

$$\begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ \hline 3.95 & 4.64 & 4.34 \\ 4.27 & 5.02 & 4.69 \\ 2.42 & 2.85 & 2.66 \\ 3.97 & 4.67 & 4.36 \\ 3.64 & 4.28 & 4.00 \\ 3.69 & 4.33 & 4.05 \\ 3.33 & 3.92 & 3.66 \\ 3.08 & 3.63 & 3.39 \\ 4.55 & 5.35 & 5.00 \end{array} = \begin{array}{|c|} 4.34 \\ 4.69 \\ 2.66 \\ 4.36 \\ 4.00 \\ 4.05 \\ 3.66 \\ 3.39 \\ 5.00 \end{array} * \begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ 0.91 & 1.07 & 1.00 \end{array}$$

which is HoleDifficulty * PlayerAbility

Notice that the HoleDifficulty factor is almost the average of that hole for the 3 players. For example hole 5, where everyone scored 4, does have a factor of 4.00. However hole 6, where the average score is also 4, has a factor of 4.05 instead of 4.00. Similarly, the PlayerAbility is almost the percentage of par that the player achieved, For example Tiger shot 39 with par being 36, and $39/36 = 1.08$ which is almost his PlayerAbility factor (for these 9 holes) of 1.07.

Why don't the hole averages and par percentages exactly match the 1-D SVD factors? The answer is that SVD further refines those numbers in a cycle. For example, we can start by assuming HoleDifficulty is the hole average and then ask what PlayerAbility best matches the scores, given those HoleDifficulty numbers? Once we have that answer we can go back and ask what HoleDifficulty best matches the scores given those PlayerAbility numbers? We keep iterating this way until we converge to a set of factors that best predict the score. SVD shortcuts this process and immediately give us the factors that we would have converged to if we carried out the process.

One very useful property of SVD is that it always finds the optimal set of factors that best predict the scores, according to the standard matrix similarity measure (the Frobenius norm). That is, if we use SVD to find the factors of a matrix, those are the best factors that can be found. This optimality property means that we don't have to wonder if a different set of numbers might predict scores better.

Now let's look at the difference between the actual scores and our 1-D approximation. A plus difference means that the actual score is higher than the predicted score, a minus difference means the actual score is lower than the prediction. For example, on the first hole Tiger got a 4 and the predicted score was 4.64 so we get $4 - 4.64 = -0.64$. In other words, we must add -0.64 to our prediction to get the actual score.

Once these differences have been found, we can do the same thing again and predict these differences using the formula HoleDifficulty2 * PlayerAbility2. Since these factors are trying to predict the differences, they are the 2-D factors and we have put a 2 after their names (ex. HoleDifficulty2) to show they are the second set of factors.

Phil	Tiger	Vijay
0.05	-0.64	0.66
-0.28	-0.02	0.31
0.58	0.15	-0.66
0.03	0.33	-0.36
0.36	-0.28	0.00
-0.69	0.67	-0.05
0.67	0.08	-0.66
-1.08	0.37	0.61
0.45	-0.35	0.00

=

-0.18
-0.38
0.80
0.15
0.35
-0.67
0.89
-1.29
0.44

*

Phil	Tiger	Vijay
0.82	-0.20	-0.53

which is HoleDifficulty(2) * PlayerAbility(2)

There are some interesting observations we can make about these factors. Notice that hole 8 has the most significant HoleDifficulty2 factor (1.29). That means that it is the hardest hole to predict. Indeed, it was the only hole on which none of the 3 players made par. It was especially hard to predict because it was the most difficult hole relative to par $(HoleDifficulty - par) = (3.39 - 3) = 0.39$, and yet Phil birdied it making his score more than a stroke below his predicted score (he scored 2 versus his predicted score of 3.08). Other holes that were hard to predict were holes 3 (0.80) and 7 (0.89) because Vijay beat Phil on those holes even though, in general, Phil was playing better.

The full SVD for this example matrix (9 holes by 3 players) has 3 sets of factors. In general, a m x n matrix where m ¿= n can have at most n factors, so our $9x3$ matrix cannot have more than 3 sets of factors. Here is the full SVD factorization (to two decimal places).

Phil	Tiger	Vijay
4	4	5
4	5	5
3	3	2
4	5	4
4	4	4
3	5	4
4	4	3
2	4	4
5	5	5

=

4.34	-0.18	-0.90
4.69	-0.38	-0.15
2.66	0.80	0.40
4.36	0.15	0.47
4.00	0.35	-0.29
4.05	-0.67	0.68
3.66	0.89	0.33
3.39	-1.29	0.14
5.00	0.44	-0.36

*

Phil	Tiger	Vijay
0.91	1.07	1.00
0.82	-0.20	-0.53
-0.21	0.76	-0.62

which is HoleDifficulty(1-3) * PlayerAbility(1-3)

By SVD convention, the HoleDifficulty and PlayerAbility vectors should all have length 1, so the conventional SVD factorization is:

$$\begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ \hline 4 & 4 & 5 \\ 4 & 5 & 5 \\ 3 & 3 & 2 \\ 4 & 5 & 4 \\ 4 & 4 & 4 \\ 3 & 5 & 4 \\ 4 & 4 & 3 \\ 2 & 4 & 4 \\ 5 & 5 & 5 \end{array} = \begin{array}{|c|c|c|} 0.35 & 0.09 & -0.64 \\ 0.38 & 0.19 & -0.10 \\ 0.22 & -0.40 & 0.28 \\ 0.36 & -0.08 & 0.33 \\ 0.33 & -0.18 & -0.20 \\ 0.33 & 0.33 & 0.48 \\ 0.30 & -0.44 & 0.23 \\ 0.28 & 0.64 & 0.10 \\ 0.41 & -0.22 & -0.25 \end{array} * \begin{array}{|c|c|c|} 21.07 & 0 & 0 \\ 0 & 2.01 & 0 \\ 0 & 0 & 1.42 \end{array} * \begin{array}{|c|c|c|} \text{Phil} & \text{Tiger} & \text{Vijay} \\ \hline 0.53 & 0.62 & 0.58 \\ -0.82 & 0.20 & 0.53 \\ -0.21 & 0.76 & -0.62 \end{array}$$

which is HoleDifficulty(1-3)* ScaleFactor(1-3) * PlayerAbility(1-3)

We hope that you have some idea of what SVD is and how it can be used. The next section covers applying SVD to Latent Sematic Analysis or LSA. Although the domain is different, the concepts are the same. We are trying to predict patterns of how words occur in documents instead of trying to predict patterns of how players score on holes.

Chapter 3

Quaternions

from[Alt05]:

> Quaternions are inextricably linked to rotations. Rotations, however, are an accident of three-dimensional space. In spaces of any other dimensions, the fundamental operations are reflections (mirrors). The quaternion algebra is, in fact, merely a sub-algebra of the Clifford algebra of order three. If the quaternion algebra might be labelled the algebra of rotations, then the Clifford algebra is the algebra of mirrors and it is thus vastly more general than quaternion algebra.

Peter Guthrie Tait, Robert S. Sutor, Timothy Daly

Preface

The Theory of Quaternions is due to Sir William Rowan Hamilton, Royal Astronomer of Ireland, who presented his first paper on the subject to the Royal Irish Academy in 1843. His Lectures on Quaternions were published in 1853, and his Elements, in 1866, shortly after his death. The Elements of Quaternions by Tait[Ta1890] is the accepted text-book for advanced students.

Large portions of this file are derived from a public domain version of Tait's book combined with the algebra available in Axiom. The purpose is to develop a tutorial introduction to the Axiom domain and its uses.

3.1 Quaternions

3.2 Vectors, and their Composition

1. For at least two centuries the geometrical representation of the negative and imaginary algebraic quantities, -1 and $\sqrt{-1}$ has been a favourite subject of speculation with mathematicians. The essence of almost all of the proposed processes consists in employing such expressions to indicate the DIRECTION, not the *length*, of lines.

2. Thus it was long ago seen that if positive quantities were measured off in one direction along a fixed line, a useful and lawful convention enabled us to express negative quantities of the same kind by simply laying them off on the same line in the opposite direction. This convention is an essential part of the Cartesian method, and is constantly employed in Analytical Geometry and Applied Mathematics.

3. Wallis, towards the end of the seventeenth century, proposed to represent the impossible roots of a quadratic equation by going *out* of the line on which, if real, they would have been laid off. This construction is equivalent to the consideration of $\sqrt{-1}$ as a directed unit-line perpendicular to that on which real quantities are measured.

4. In the usual notation of Analytical Geometry of two dimensions, when rectangular axes are employed, this amounts to reckoning each unit of length along Oy as $+\sqrt{-1}$, and on Oy' as $-\sqrt{-1}$; while on Ox each unit is $+1$, and on Ox it is -1.

If we look at these four lines in circular order, i.e. in the order of positive rotation (that of the northern hemisphere of the earth about its axis, or *opposite* to that of the hands of a watch), they give

$$1, \sqrt{-1}, -1, -\sqrt{-1}$$

In Axiom the same elements would be written as complex numbers which are constructed using the function **complex**:

```
complex(1,0)
```

$$1$$

Type: Complex Integer

```
complex(0,1)
```

$$\%i$$

Type: Complex Integer

```
complex(-1,0)
```

$$-1$$

Type: Complex Integer

```
complex(0,-1)
```

$$-i$$

Type: Complex Integer

Note that %i is of type Complex(Integer), that is, the imaginary part of a complex number. The apparently equivalent expression

```
sqrt(-1)
```

$$\sqrt{-1}$$

Type: AlgebraicNumber

has the type AlgebraicNumber which means that it is the root of a polynomial with rational coefficients.

In this series each expression is derived from that which precedes it by multiplication by the factor $\sqrt{-1}$. Hence we may consider $\sqrt{-1}$ as an operator, analogous to a handle perpendicular to the plane of xy, whose effect on any line in that plane is to make it rotate (positively) about the origin through an angle of 90°.

In Axiom

```
%i*%i
```

$$-1$$

Type: Complex Integer

5. In such a system, (which seems to have been first developed, in 1805, by Buée) a point in the plane of reference is defined by a single imaginary expression. Thus $a + b\sqrt{-1}$ may be considered as a single quantity, denoting the point, P, whose coordinates are a and b. Or, it may be used as an expression for the line OP joining that point with the origin. In the latter sense, the expression $a + b\sqrt{-1}$ implicitly contains the *direction*, as well as the *length*, of this line ; since, as we see at once, the direction is inclined at an angle $\tan^{-1}(b/a)$ to the axis of x, and the length is $\sqrt{a^2+b^2}$. Thus, say we have

$$OP = a + b\sqrt{-1}$$

the line OP considered as that by which we pass from one extremity, O, to the other, P. In this sense it is called a VECTOR. Considering, in the plane, any other vector,

$$OQ = a^{'} + b^{'}\sqrt{-1}$$

In order to created superscripted variables we use the superscript function from the SYMBOL domain. So we can create $a^{'}$ as "ap" (that is, "a-prime") and $b^{'}$ as "bp" ("b-prime") thus (also note that the underscore character is Axiom's escape character which removes any special meaning of the next character, in this case, the quote character):

```
ap:=superscript(a,[_'])
```

$$a^{'}$$

Type: Symbol

```
bp:=superscript(b,[_'])
```

$$b^{'}$$

Type: Symbol

at this point we can type

```
ap+bp*%i
```

$$a^{'} + b^{'}\ \%i$$

Type: Complex Polynomial Integer

the addition of these two lines obviously gives

$$OR = a + a^{'} + (b + b^{'})\sqrt{-1}$$

In Axiom the computation looks like:

```
op:=complex(a,b)
```

$$a + b\ \%i$$

Type: Complex Polynomial Integer

```
oq:=complex(ap,bp)
```

$$a' + b' \%i$$

Type: Complex Polynomial Integer

```
op + oq
```

$$a + a' + (b + b')\%i$$

Type: Complex Polynomial Integer

and we see that the sum is the diagonal of the parallelogram on OP, OQ. This is the law of the composition of simultaneous velocities; and it contains, of course, the law of subtraction of one directed line from another.

6. Operating on the first of these symbols by the factor $\sqrt{-1}$, it becomes $-b + a\sqrt{-1}$; and now, of course, denotes the point whose x and y coordinates are $-b$ and a; or the line joining this point with the origin. The length is still $\sqrt{a^2 + b^2}$, but the angle the line makes with the axis of x is $\tan^{-1}(-a/b)$; which is evidently greater by $\pi/2$ than before the operation.

```
op*complex(0,1)
```

$$-b + a\ i$$

Type: Complex Polynomial Integer

7. De Moivre's Theorem tends to lead us still further in the same direction. In fact, it is easy to see that if we use, instead of $\sqrt{-1}$, the more general factor $\cos\alpha + \sqrt{-1}\sin\alpha$, its effect on any line is to turn it through the (positive) angle α. in the plane of x, y. [Of course the former factor, $\sqrt{-1}$, is merely the particular case of this, when $\alpha = \frac{\pi}{2}$].

Thus

$$\begin{array}{rl} & (\cos\alpha + \sqrt{-1}\sin\alpha)(a + b\sqrt{-1}) \\ = & a\cos\alpha - b\sin\alpha + \sqrt{-1}(a\sin\alpha + b\cos\alpha) \end{array}$$

by direct multiplication. The reader will at once see that the new form indicates that a rotation through an angle α has taken place, if he compares it with the common formulae for turning the coordinate axes through a given angle. Or, in a less simple manner, thus

$$\begin{array}{rcl} Length & = & \sqrt{(a\cos\alpha - b\sin\alpha)^2 + (a\sin\alpha + b\cos\alpha)^2} \\ & = & \sqrt{a^2 + b^2} \end{array}$$

as before.

Inclination to axis of x

$$\begin{aligned} &= \tan^{-1}\tfrac{a\sin\alpha+b\cos\alpha}{a\cos\alpha-b\sin\alpha} \\ &= \tan^{-1}\tfrac{\tan\alpha+\frac{b}{a}}{1-\frac{b}{a}\tan\alpha} \\ &= \alpha+\tan^{-1}\tfrac{b}{a} \end{aligned}$$

8. We see now, as it were, why it happens that

$$(\cos\alpha+\sqrt{-1}\sin\alpha)^m = \cos m\alpha+\sqrt{-1}\sin m\alpha$$

In fact, the first operator produces m successive rotations in the same direction, each through the angle α ; the second, a single rotation through the angle $m\alpha$.

9. It may be interesting, at this stage, to anticipate so far as to remark that in the theory of Quaternions the analogue of

$$\begin{array}{ll} & \cos\theta+\sqrt{-1}\sin\theta \\ \text{is} & \cos\theta+\omega\sin\theta \\ \text{where} & \omega^2=-1 \end{array}$$

Here, however, ω is not the algebraic $\sqrt{-1}$, but is *any directed unit-line* whatever in space.

10. In the present century Argand, Warren, Mourey, and others, extended the results of Wallis and Buée. They attempted to express as a line the product of two lines each represented by a symbol such $a+b\sqrt{-1}$. To a certain extent they succeeded, but all their results remained confined to two dimensions.

The product, $\prod$, of two such lines was defined as the fourth proportional to unity and the two lines, thus

$$\begin{array}{ll} & 1 : a+b\sqrt{-1} :: a^{'}+b^{'}\sqrt{-1} : \prod \\ \text{or} & \prod = (aa^{'}-bb^{'})+(a^{'}b+b^{'}a)\sqrt{-1} \end{array}$$

The length of $\prod$ is obviously the product of the lengths of the factor lines; and its direction makes an angle with the axis of x which is the sum of those made by the factor lines. From this result the quotient of two such lines follows immediately.

11. A very curious speculation, due to Servois and published in 1813 in Gergonne's *Annales*, is one of the very few, so far as has been discovered, in which a well-founded guess at a possible mode of extension to three dimensions is contained. Endeavouring to extend to *space* the form $a+b\sqrt{-1}$ for the plane, he is guided by analogy to write for a directed unit-line in space the form

$$p\cos\alpha+q\cos\beta+r\cos\gamma$$

where α, β, γ are its inclinations to the three axes. He perceives easily that p, q, r must be *non-reals*: but, he asks, “seraient-elles *imaginaires* réductibles à la forme générale $A + B\sqrt{-1}$?” The i,j, k of the Quaternion Calculus furnish an answer to this question. (See Chap. II.) But it may be remarked that, in applying the idea to lines in a plane, a vector OP will no longer be represented (as in §5) by

$$\begin{array}{lrcr} & \mathrm{OP} & = & a+b\sqrt{-1} \\ \text{but by} & \mathrm{OP} & = & pa+qb \\ \text{And if, similarly,} & \mathrm{OQ} & = & pa^{'}+qb^{'} \end{array}$$

the addition of these two lines gives for OR (which retains its previous signification)

$$OR = p(a + a^{'} + q(b + b^{'})$$

12. Beyond this, few attempts were made, or at least recorded, in earlier times, to extend the principle to space of three dimensions; and, though many such had been made before 1843, none, with the single exception of Hamilton’s, have resulted in simple, practical methods; all, however ingenious, seeming to lead almost at once to processes and results of fearful complexity.

For a lucid, complete, and most impartial statement of the claims of his predecessors in this field we refer to the Preface to Hamilton’s *Lectures on Quaternions*. He there shows how his long protracted investigations of Sets culminated in this unique system of tridimensional-space geometry.

13. It was reserved for Hamilton to discover the use and properties of a class of symbols which, though all in a certain sense square roots of -1, may be considered as *real* unit lines, tied down to no particular direction in space ; the expression for a vector is, or may be taken to be,

$$\rho = ix + jy + kz$$

but such vector is considered in connection with an *extraspatial* magnitude w, and we have thus the notion of a QUATERNION

$$w + \rho$$

This is the fundamental notion in the singularly elegant, and enormously powerful, Calculus of Quaternions.

While the schemes for using the algebraic $\sqrt{-1}$ to indicate direction make one direction in space expressible by real numbers, the remainder being imaginaries of some kind, and thus lead to expressions which are heterogeneous ; Hamilton s system makes all directions in space equally imaginary, or rather equally real, thereby ensuring to his Calculus the power of dealing with space indifferently in all directions.

In fact, as we shall see, the Quaternion method is independent of axes or any supposed directions in space, and takes its reference lines solely from the problem it is applied to.

14. But, for the purpose of elementary exposition, it is best to begin by assimilating it as closely as we can to the ordinary Cartesian methods of Geometry of Three Dimensions, with which the student is supposed to be, to some extent at least, acquainted. Such assistance, it will be found, can (as a rule) soon be dispensed with; and Hamilton regarded any apparent necessity for an oc casional recurrence to it, in higher applications, as an indication of imperfect development in the proper methods of the new Calculus.

We commence, therefore, with some very elementary geometrical ideas, relating to the theory of vectors in space. It will subsequently appear how we are thus led to the notion of a Quaternion.

15. Suppose we have two points A and B in *space*, and suppose A given, on how many numbers does B's relative position depend ?

If we refer to Cartesian coordinates (rectangular or not) we find that the data required are the excesses of B's three coordinates over those of A. Hence three numbers are required.

Or we may take polar coordinates. To define the moon's position with respect to the earth we must have its Geocentric Latitude and Longitude, or its Right Ascension and Declination, and, in addition, its distance or radius-vector. *Three* again.

16. Here it is to be carefully noticed that nothing has been said of the *actual* coordinates of either A or B, or of the earth and moon, in space; it is only the *relative* coordinates that are contemplated.

Hence any expression, as $\overline{AB}$, denoting a line considered with reference to direction and currency as well as length, (whatever may be its actual position in space) contains implicitly *three* numbers, and all lines parallel and equal to AB, and concurrent with it, depend in the same way upon the same three. Hence, *all lines which are equal, parallel, and concurrent, may be represented by a common symbol, and that symbol contains three distinct numbers.* In this sense a line is called a VECTOR, since by it we pass from the one extremity, A, to the other, B, and it may thus be considered as an instrument which *carries* A to B: so that a vector may be employed to indicate a definite *translation* in space.

[The term " currency " has been suggested by Cayley for use instead of the somewhat vague suggestion sometimes taken to be involved in the word "direction." Thus parallel lines have the same direction, though they may have similar or opposite currencies. The definition of a vector essentially includes its currency.]

17. We may here remark, once for all, that in establishing a new Calculus, we are at liberty to give any definitions whatever of our symbols, provided that no two of these interfere with, or contradict, each other, and in doing so in

Quaternions sl simplicity and (so to speak) *naturalness* were the inventor's aim.

18. Let $\overline{AB}$ be represented by α, we know that α involves *three* separate numbers, and that these depend solely upon the position of B *relatively* to A. Now if CD be equal in length to AB and if these lines be parallel, and have the same currency, we may evidently write

$$\overline{CD} = \overline{AB} = \alpha$$

where it will be seen that the sign of equality between vectors contains implicitly *equality in length*, *parallelism in direction*, and *concurrency*. So far we have *extended* the meaning of an algebraical symbol. And it is to be noticed that an equation between vectors, as

$$\alpha = \beta$$

contains *three* distinct equations between mere numbers.

19. We must now define + (and the meaning of − will follow) in the new Calculus. Let A, B, C be any three points, and (with the above meaning of =) let

$$\overline{AB} = \alpha, \overline{BC} = \beta, \overline{AC} = \gamma$$

If we define + (in accordance with the idea (§16) that a vector represents a *translation*) by the equation

$$\alpha + \beta = \gamma$$

$$\text{or} \qquad \overline{AB} + \overline{BC} = \overline{AC}$$

we contradict nothing that precedes, but we at once introduce the idea that *vectors are to be compounded*, *in direction and magnitude*, *like simultaneous velocities*. A reason for this may be seen in another way if we remember that by *adding* the (algebraic) differences of the Cartesian coordinates of B and A, to those of the coordinates of C and B, we get those of the coordinates of C and A. Hence these coordinates enter *linearly* into the expression for a vector. (See, again, §5.)

20. But we also see that if C and A coincide (and C may be *any* point)

$$\overline{AC} = 0$$

for no vector is then required to carry A to C. Hence the above relation may be written, in this case,

$$\overline{AB} + \overline{BA} = 0$$

or, introducing, and by the same act defining, the symbol −,

$$\overline{AB} = -\overline{BA}$$

Hence, *the symbol −, applied to a vector, simply shows that its currency is to be reversed*. And this is consistent with all that precedes; for instance,

$$\begin{array}{lrcl} & \overline{AB} + \overline{BC} & = & \overline{AC} \\ \text{and} & \overline{AB} = \overline{AC} & - & \overline{BC} \\ \text{or} & = \overline{AC} & + & \overline{CB} \end{array}$$

are evidently but different expressions of the same truth.

21. In any triangle, ABC, we have, of course,

$$\overline{AB} + \overline{BC} + \overline{CA} = 0$$

and, in any closed polygon, whether plane or gauche,

$$\overline{AB} + \overline{BC} + \ldots + \overline{YZ} + \overline{ZA} = 0$$

In the case of the polygon we have also

$$\overline{AB} + \overline{BC} + \ldots + \overline{YZ} = \overline{AZ}$$

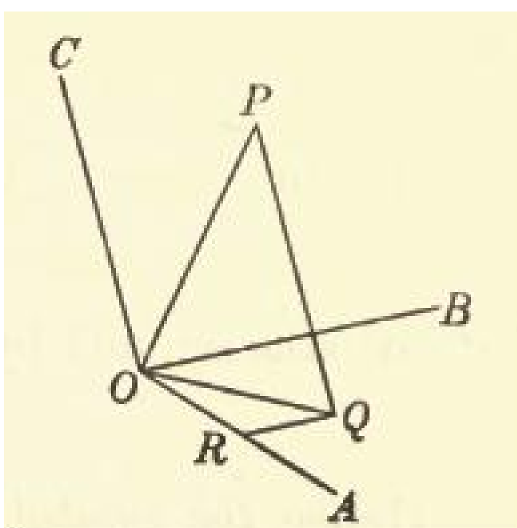
These are the well-known propositions regarding composition of velocities, which, by Newton's second law of motion, give us the geometrical laws of composition of forces acting at one point.

22. If we compound any number of **parallel** vectors, the result is obviously a numerical multiple of any one of them. Thus, if A, B, C are in one straight line,

$$\overline{BC} = x\overline{AB}$$

where x is a number, positive when B lies between A and C, otherwise negative; but such that its numerical value, independent of sign, is the ratio of the length of BC to that of AB. This is at once evident if AB and BC be commensurable; and is easily extended to incommensurables by the usual *reductio ad absurdum*.

23. An important, but almost obvious, proposition is that *any vector may be resolved, and in one way only, into three components parallel respectively to any three given vectors, no two of which are parallel, and which are not parallel to one plane.*



Let OA, OB, OC be the three fixed vectors, OP any other vector. From P draw PQ parallel to CO, meeting the plane BOA in Q. [There must be a definite point Q, else PQ, and therefore CO, would be parallel to BOA, a case specially excepted.] From Q draw QR parallel to BO, meeting OA in R.

Then we have $\overline{OP} = \overline{OR} + \overline{RQ} + \overline{QP}$ (§21), and these components are respectively parallel to the three given vectors. By §22 we may express $\overline{OR}$ as a numerical multiple of $\overline{OA}$, $\overline{RQ}$ of $\overline{OB}$, and $\overline{QP}$ of $\overline{OC}$. Hence we have, generally, for any vector in terms of three fixed non-coplanar vectors, α, β, γ

$$\overline{OP} = \rho = x\alpha + y\beta + z\gamma$$

which exhibits, in one form, the *three* numbers on which a vector depends (§16). Here x, y, z are perfectly definite, and can have but single values.

24. Similarly any vector, as $\overline{OQ}$, in the same plane with $\overline{OA}$ and $\overline{OB}$, can be resolved (in one way only) into components $\overline{OR}$, $\overline{RQ}$, parallel respectively to $\overline{OA}$ and $\overline{OB}$; so long, at least, as these two vectors are not parallel to each other.

25. There is particular advantage, in certain cases, in employing a series of *three mutually perpendicular unit-vectors* as lines of reference. This system Hamilton denotes by i,j, k.

Any other vector is then expressible as

$$\rho = xi + yj + zk$$

Since i, j, k are unit-vectors, x, y, z are here the lengths of conterminous edges of a rectangular parallelepiped of which ρ is the vector-diagonal; so that the length of ρ is, in this case,

$$\sqrt{x^2 + y^2 + z^2}$$

Let $$\omega = \xi i + \eta j + \zeta k$$
be any other vector, then (by the proposition of §23) the vector

equation $\qquad \rho = \omega$

obviously involves the following three equations among numbers,

$$x = \xi, y = \eta, z = \zeta$$

Suppose i to be drawn eastwards, j northwards, and k upwards, this is equivalent merely to saying that *if two points coincide, they are equally to the east (or west) of any third point, equally to the north (or south) of it, and equally elevated above (or depressed below) its level.*

26. It is to be carefully noticed that it is only when α, β, γ are not coplanar that a vector equation such as

$$\rho = \omega$$

or
$$x\alpha + y\beta + z\gamma = \xi\alpha + \eta\beta + \zeta\gamma$$
necessitates the three numerical equations

$$x = \xi, y = \eta, z = \zeta$$

For, if α, β, γ be coplanar (§24), a condition of the following form must hold

$$\gamma = a\alpha + b\beta$$

Hence,
$$\rho = (x + za)\alpha + (y + zb)\beta$$
$$\omega = (\xi + \zeta a)\alpha + (\eta + \zeta b)\beta$$
and the equation
$$\rho = \omega$$
now requires only the two numerical conditions

$$x + za = \xi + \zeta a \qquad y + zb = \eta + \zeta b$$

27. *The Commutative and Associative Laws hold in the combination of vectors by the signs + and −*. It is obvious that, if we prove this for the sign +, it will be equally proved for −, because − before a vector (§20) merely indicates that it is to be reversed before being considered positive.

Let A, B, C, D be, in order, the corners of a parallelogram ; we have, obviously,

$$\overline{AB} = \overline{DC} \qquad \overline{AD} = \overline{BC}$$

And
$$\overline{AB} + \overline{BC} = \overline{AC} = \overline{AD} + \overline{DC} = \overline{BC} + \overline{AB}$$

Hence the commutative law is true for the addition of any two vectors, and is therefore generally true.

Again, whatever four points are represented by A, B, C, D, we

$$\overline{AD} = \overline{AB} + \overline{BD} = \overline{AC} + \overline{CD}$$

or substituting their values for $\overline{AD}$, $\overline{BD}$, $\overline{AC}$ respectively, in these three expressions,

$$\overline{AB} + \overline{BC} + \overline{CD} = \overline{AB} + (\overline{BC} + \overline{CD}) = (\overline{AB} + \overline{BC}) + \overline{CD}$$

And thus the truth of the associative law is evident.

28. The equation
$$\rho = x\beta$$

where ρ is the vector connecting a variable point with the origin, β a definite vector, and x an indefinite number, represents the straight line drawn from the origin parallel to β (§22).

The straight line drawn from A, where $\overline{OA} = \alpha$, and parallel to β, has the equation
$$\rho = \alpha + x\beta \tag{1}$$

In words, we may pass directly from O to P by the vector $\overline{OP}$ or ρ; or we may pass first to A, by means of $\overline{OA}$ or α, and then to P along a vector parallel to β (§16).

Equation (1) is one of the many useful forms into which Quaternions enable us to throw the general equation of a straight line in space. As we have seen (§25) it is equivalent to three numerical equations; but, as these involve the indefinite quantity x, they are virtually equivalent to but *two*, as in ordinary Geometry of Three Dimensions.

29. A good illustration of this remark is furnished by the fact that the equation

$$\rho = y\alpha + x\beta$$

which contains two indefinite quantities, is virtually equivalent to only one numerical equation. And it is easy to see that it represents the plane in which the lines α and β lie; or the surface which is formed by drawing, through every point of OA, a line parallel to OB. In fact, the equation, as written, is simply §24 in symbols.

And it is evident that the equation

$$\rho = \gamma + y\alpha + x\beta$$

is the equation of the plane passing through the extremity of γ, and parallel to α and β.

It will now be obvious to the reader that the equation

$$\rho = p_1\alpha_1 + p_2\alpha_2 + \ldots = \sum p\alpha$$

where α_1, α_2 , &c. are given vectors, and p_1, p_2, &c. numerical quantities, *represents a straight line* if p_1, p_2, &c. be linear functions of *one* indeterminate number; and a *plane*, if they be linear expressions containing two indeterminate numbers. Later (§31 (l)), this theorem will be much extended.

Again, the equation

$$\rho = x\alpha + y\beta + z\gamma$$

refers to *any* point whatever in space, provided α, β, γ are not coplanar. (Ante, §23)

30. The equation of the line joining any two points A and B, where $\overline{OA} = \alpha$ and $\overline{OB} = \beta$, is obviously

$$\rho = \alpha + x(\beta - \alpha)$$

or

$$\rho = \beta + y(\alpha - \beta)$$

These equations are of course identical, as may be seen by putting $1 - y$ for x.

The first may be written

$$\rho + (x - 1)\alpha - x\beta = 0$$

or $$p\rho + q\alpha + r\beta = 0$$
subject to the condition $p+q+r = 0$ identically. That is – A homogeneous linear function of three vectors, equated to zero, expresses that the extremities of these vectors are in one straight line, *if the sum of the coefficients be identically zero.*

Similarly, the equation of the plane containing the extremities A, B, C of the three non-coplanar vectors α, β, γ is

$$\rho = \alpha + x(\beta - \alpha) + y(\gamma - \beta)$$

where x and y are each indeterminate.

This may be written
$$p\rho + q\alpha + r\beta + s\gamma = 0$$
with the identical relation
$$p + q + r + x = 0$$
which is one form of the condition that four points may lie in one plane.

31. We have already the means of proving, in a very simple manner, numerous classes of propositions in plane and solid geometry. A very few examples, however, must suffice at this stage; since we have hardly, as yet, crossed the threshold of the subject, and are dealing with mere linear equations connecting two or more vectors, and even with them *we are restricted as yet to operations of mere addition.* We will give these examples with a painful minuteness of detail, which the reader will soon find to be necessary only for a short time, if at all.

(a) *The diagonals of a parallelogram bisect each other.*

Let $ABCD$ be the parallelogram, O the point of intersection of its diagonals. Then
$$\overline{AO} + \overline{OB} = \overline{AB} = \overline{DC} = \overline{DO} + \overline{OC}$$
which gives $$\overline{AO} - \overline{OC} = \overline{DO} - \overline{OB}$$
The two vectors here equated are parallel to the diagonals respectively. Such an equation is, of course, absurd unless

1. The diagonals are parallel, in which case the figure is not a parallelogram;
2. $\overline{AO} = \overline{OC}$, and $\overline{DO} = \overline{OB}$, the proposition.

(b) *To shew that a triangle can be constructed, whose sides are parallel, and equal, to the bisectors of the sides of any triangle.*

Let ABC be any triangle, Aa, Bb, Cc the bisectors of the sides.

Then
$$\begin{array}{lll} \overline{Aa} & = \overline{AB} + \overline{Ba} & = \overline{AB} + \frac{1}{2}\overline{BC} \\ \overline{Bb} & \quad\ldots & = \overline{BC} + \frac{1}{2}\overline{CA} \\ \overline{Cc} & \quad\ldots & = \overline{CA} + \frac{1}{2}\overline{AB} \end{array}$$

Hence $$\overline{Aa} + \overline{Bb} + \overline{Cc} = \tfrac{3}{2}(\overline{AB} + \overline{BC} + \overline{CA}) = 0$$
which (§21) proves the proposition.

Also
$$\begin{aligned}\overline{Aa} &= \overline{AB} + \tfrac{1}{2}\overline{BC}\\ &= \overline{AB} - \tfrac{1}{2}(\overline{CA} + \overline{AB})\\ &= \tfrac{1}{2}(\overline{AB} - \overline{CA})\\ &= \tfrac{1}{2}(\overline{AB} + \overline{AC})\end{aligned}$$

results which are sometimes useful. They may be easily verified by producing $\overline{Aa}$ to twice its length and joining the extremity with B.

(b') *The bisectors of the sides of a triangle meet in a point, which trisects each of them.*

Taking A as origin, and putting α, β, γ for vectors parallel, and equal, to the sides taken in order BC, CA, AB; the equation of Bb is (§28 (1))

$$\rho = \gamma + x(\gamma + \frac{\beta}{2}) = (1+x)\gamma + \frac{x}{2}\beta$$

That of Cc is, in the same way,

$$\rho = -(1+y)\beta - \frac{y}{2}\gamma$$

At the point O, where Bb and Cc intersect,

$$\rho = (1+x)\gamma + \frac{x}{2}\beta = -(1+y)\beta - \frac{y}{2}\gamma$$

Since γ and β are not parallel, this equation gives

$$1 + x = -\frac{y}{2} \text{ and } \frac{x}{2} = -(1+y)$$

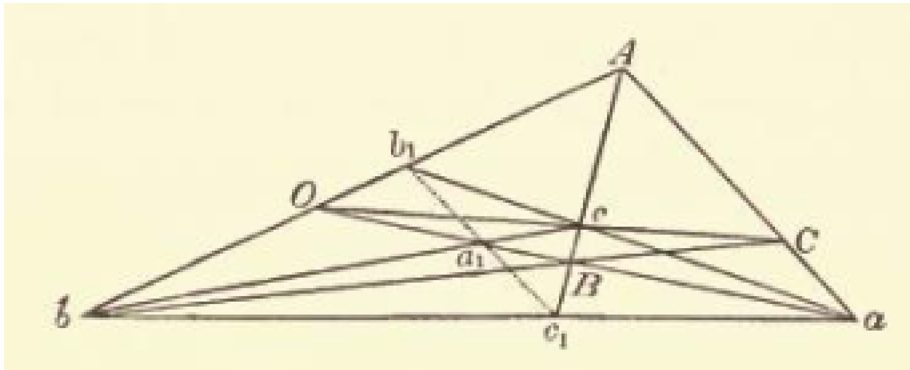
From these $$x = y = -\tfrac{2}{3}$$

Hence $$\overline{AO} = \tfrac{1}{3}(\gamma - \beta) = \tfrac{2}{3}\overline{Aa} \text{ (See Ex. (b))}$$

This equation shows, being a vector one, that $\overline{Aa}$ passes through O, and that $AO : Oa :: 2{:}1$.

(c) If
$$\overline{OA} = \alpha$$
$$\overline{OB} = \beta$$
$$\overline{OC} = l\alpha + m\beta$$

be three given co-planar vectors, c the intersection of AB, OC, and if the lines indicated in the figure be drawn, the points a_1,b_1,c_1 lie in a straight line.



We see at once, by the process indicated in §30, that

$$\overline{Oc} = \frac{l\alpha + m\beta}{l + m}, \qquad \overline{Ob} = \frac{l\alpha}{1 - m}, \qquad \overline{Oa} = \frac{m\beta}{1 - l}$$

Hence we easily find

$$\overline{Oa_1} = -\frac{m\beta}{1 - l - 2m}, \quad \overline{Ob_1} = -\frac{l\alpha}{1 - 2l - m}, \quad \overline{Oc_1} = \frac{-l\alpha + m\beta}{m - l}$$

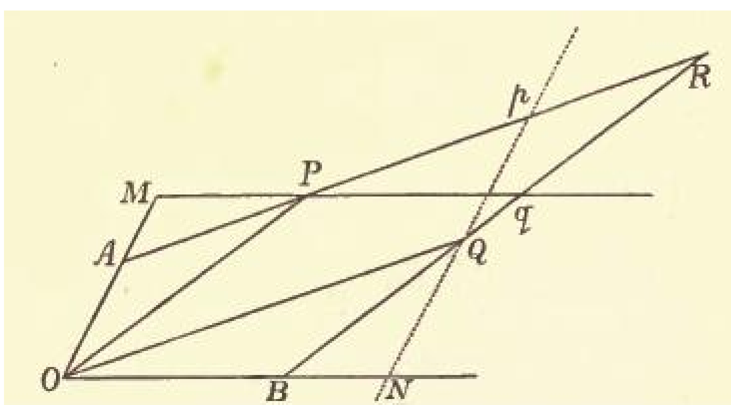
These give

$$-(1 - l - 2m)\overline{Oa_1} + (1 - 2l - m)\overline{Ob_1} - (m - l)\overline{Oc_1} = 0$$

But $\quad -(1 - l - 2m) + (1 - 2l - m) - (m - l) = 0$ identically.

This, by §30, proves the proposition.

(d) Let $\overline{OA} = \alpha$, $\overline{OB} = \beta$, be any two vectors. If MP be a given line parallel to OB; and OQ, BQ, be drawn parallel to AP, OP respectively ; the locus of Q is a straight line parallel to OA.



Let $\qquad \overline{OM} = e\alpha$

Then $\qquad \overline{AP} = \overline{e - 1}\alpha + x\beta$

Hence the equation of OQ is

$$\rho = y(\overline{e-1}\alpha + x\beta)$$

and that of BQ is $\rho = \beta + z(e\alpha + x\beta)$
At Q we have, therefore,

$$\left.\begin{array}{r} xy = 1 + zx \\ y(e-1) = ze \end{array}\right\}$$

These give $xy = e$, and the equation of the locus of Q is

$$\rho = e\beta + y^{'}\alpha$$

i.e. a straight line parallel to OA, drawn through N in OB produced, so that

$$ON : OB :: OM : OA$$

COR. If BQ meet MP in q, $\overline{Pq} = \beta$; and if AP meet NQ in p, $\overline{Qp} = \alpha$.

Also, for the point R we have $\overline{pR} = \overline{AP}$, $\overline{QR} = \overline{Bq}$.

Further, the locus of R is a hyperbola, of which MP and NQ are the asymptotes. See, in this connection, §31 (k) below.

Hence, *if from any two points, A and B, lines be drawn intercepting a given length Pq on a given line Mq ; and if, from R their point of intersection, Rp be laid off $= PA$, and $RQ = qB$; Q and p lie on a fixed straight line, and the length of Qp is constant.*

(e) *To find the centre of inertia of any system of masses.*

If $\overline{OA} = \alpha$, $\overline{OB} = \alpha_1$, be the vector sides of any triangle, the vector from the vertex dividing the base AB in C so that

$$BC : CA :: m : m_1$$

is $\frac{m\alpha+m_1\alpha_1}{m+m_1}$

For AB is $\alpha_1 - \alpha$, and therefore $\overline{AC}$ is

$$\frac{m_1}{m+m_1}(\alpha_1 - \alpha)$$

Hence

$$\begin{aligned} \overline{OC} &= \overline{OA} + \overline{AC} \\ &= \alpha + \frac{m_1}{m+m_1}(\alpha_1 - \alpha) \\ &= \frac{m\alpha + m_1\alpha_1}{m+m_1} \end{aligned}$$

This expression shows how to find the centre of inertia of two masses ; m at the extremity of α, m_1 at that of α_1. Introduce m_2 at the extremity of a_2, then the vector of the centre of inertia of the three is, by a second application of the formula,

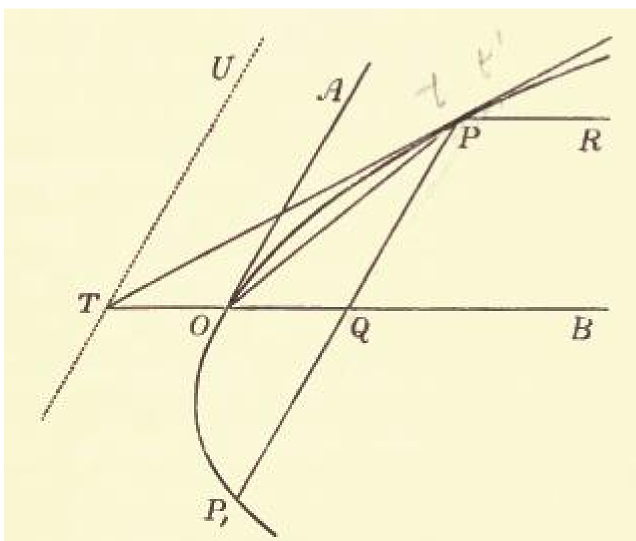
$$\frac{(m+m_1)(\frac{m\alpha+m_1\alpha_1}{m+m_1})+m_2\alpha_2}{(m+m_1)+m_2}=\frac{m\alpha+m_1\alpha_1+m_2\alpha_2}{m+m_1+m_2}$$

From this it is clear that, for any number of masses, expressed generally by m at the extremity of the vector α, the vector of the centre of inertia is

$$\beta=\frac{\sum(m\alpha)}{\sum(m)}$$

This may be written $\qquad \sum m(\alpha-\beta)=0$
Now a $\alpha_1-\beta$ is the vector of m_1 with respect to the centre of inertia. Hence the theorem, *If the vector of each element of a mass, drawn from the centre of inertia, be increased in length in proportion to the mass of the element, the sum of all these vectors is zero.*

(f) We see at once that the equation



$$\rho=\alpha t+\frac{\beta t^2}{2}$$

where t is an indeterminate number, and α, β given vectors, represents a parabola. The origin, O, is a point on the curve, β is parallel to the axis, i.e. is the diameter OB drawn from the origin, and α is OA the tangent at the origin. In the figure

$$\overline{QP}=\alpha t, \qquad \overline{OQ}=\frac{\beta t^2}{2}$$

The secant joining the points where t has the values t and $t^{'}$ is represented by the equation

$$\begin{aligned}\rho &= \alpha t + \tfrac{\beta t^2}{2} + x\left(\alpha t^{'} + \tfrac{\beta t^{'2}}{2} - \alpha t - \tfrac{\beta t^2}{2}\right) \qquad (\S 30)\\ &= \alpha t + \tfrac{\beta t^2}{2} + x(t^{'} - t)\left\{\alpha + \beta \tfrac{t^{'} - t}{2}\right\}\end{aligned}$$

Write x for $x(t^{'} - t)$ [which may have any value], then put $t^{'} = t$, and the equation of the tangent at the point (t) is

$$\rho = \alpha t + \frac{\beta t^2}{2} + x(\alpha + \beta t)$$

In this put $x = -t$, and we have

$$\rho = -\frac{\beta t^2}{2}$$

or the intercept of the tangent on the diameter is equal in length to the abscissa of the point of contact, but has the opposite currency.

Otherwise: the tangent is parallel to the vector $\alpha + \beta t$ or $\alpha t + \beta t^2$ or $\frac{\beta t^2}{2} + \alpha t + \frac{\beta t^2}{2}$ or $\overline{OQ} + \overline{OP}$. But $\overline{TP} = \overline{TO} + \overline{OP}$, hence $\overline{TO} = \overline{OQ}$.

(g) Since the equation of any tangent to the parabola is

$$\rho = \alpha t + \frac{\beta t^2}{2} + x(\alpha + \beta t)$$

let us find the tangents which can be drawn from a given point. Let the vector of the point be

$$\rho = p\alpha + q\beta \qquad (\S 24)$$

Since the tangent is to pass through this point, we have, as con ditions to determine t and x,

$$t + x = p$$

$$\frac{t^2}{2} + xt = q$$

by equating respectively the coefficients of α and β.

Hence $$t = p \pm \sqrt{p^2 - 2q}$$

Thus, in general, *two* tangents can be drawn from a given point. These coincide if

$$p^2 = 2q$$

that is, if the vector of the point from which they are to be drawn is

$$\rho = p\alpha + q\beta = p\alpha + \frac{p^2}{2}\beta$$

i.e. if the point lies on the parabola. They are imaginary if $2q > p^2$, that is, if the point be

$$\rho = p\alpha + \left(\frac{p^2}{2} + r\right)\beta$$

r being *positive*. Such a point is evidently *within* the curve, as at R, where $\overline{OQ} = \frac{p^2}{2}\beta$, $\overline{QP} = p\alpha$, $\overline{PR} = r\beta$.

(h) Calling the values of t for the two tangents found in (g) t_1 and t_2 respectively, it is obvious that the vector joining the points of contact is

$$\alpha t_1 + \frac{\beta t_1^2}{2} - \alpha t_2 - \frac{\beta t_2^2}{2}$$

which is parallel to $\alpha + \beta\frac{t_1+t_2}{2}$ or, by the values of t_1 and t_2 in (g),

$$\alpha + p\beta$$

Its direction, therefore, does not depend on q. In words, *If pairs of tangents be drawn to a parabola from points of a diameter produced, the chords of contact are parallel to the tangent at the vertex of the diameter.* This is also proved by a former result, for we must have $\overline{OT}$ for each tangent equal to $\overline{QO}$.

(i) The equation of the chord of contact, for the point whose vector is

$$\rho = p\alpha + q\beta$$

is thus $$\rho = \alpha t_1 + \frac{\beta t_1^2}{2} + y(\alpha + p\beta)$$

Suppose this to pass always through the point whose vector is

$$\rho = a\alpha + b\beta$$

Then we must have

$$\left.\begin{array}{rcl} t_1 + y &=& a \\ \frac{t_1^2}{2} + py &=& b \end{array}\right\}$$

or $$t_1 = p \pm \sqrt{p^2 - 2p\alpha + 2\beta}$$

Comparing this with the expression in (g), we have

$$q = pa - b$$

that is, the point from which the tangents are drawn has the vector a straight line (§28 (1)).

The mere form of this expression contains the proof of the usual properties of the pole and polar in the parabola ; but, for the sake of the beginner, we adopt a simpler, though equally general, process.

Suppose $\alpha = 0$. This merely restricts the pole to the particular diameter to which we have referred the parabola. Then the pole is Q, where

$$\rho = b\beta$$

and the polar is the line TU, for which

$$\rho = -b\beta + p\alpha$$

Hence the polar of any point is parallel to the tangent at the extremity of the diameter on which the point lies, and its intersection with that diameter is as far beyond the vertex as the pole is within, and vice versa.

(j) As another example let us prove the following theorem. *If a triangle be inscribed in a parabola, the three points in which the sides are met by tangents at the angles lie in a straight line.*

Since O is any point of the curve, we may take it as one corner of the triangle. Let t and t_1 determine the others. Then, if ω_1,ω_2,ω_3 represent the vectors of the points of intersection of the tangents with the sides, we easily find

$$\begin{aligned} \omega_1 &= \tfrac{t_1^2}{2t_1-t}\left(\alpha + \tfrac{t}{2}\beta\right) \\ \omega_2 &= \tfrac{t^2}{2t-t_1}\left(\alpha + \tfrac{t_1}{2}\beta\right) \\ \omega_3 &= \tfrac{tt_1}{t_1+t}\alpha \end{aligned}$$

These values give

$$\frac{2t_1-t}{t_1}\omega_1 - \frac{2t-t_1}{t}\omega_2 - \frac{t_1^2-t^2}{tt_1}\omega_3 = 0$$

Also

$$\frac{2t_1-t}{t_1} - \frac{2t-t_1}{t} - \frac{t_1^2-t^2}{tt_1} = 0$$

identically.

Hence, by §30, the proposition is proved.

(k) Other interesting examples of this method of treating curves will, of course, suggest themselves to the student. Thus

$$\rho = \alpha\cos t + \beta\sin t$$

or

$$\rho = \alpha x + \beta\sqrt{1-x^2}$$

represents an ellipse, of which the given vectors α and β are semiconjugate diameters. If t represent time, the radius-vector of this ellipse traces out equal areas in equal times. [We may anticipate so far as to write the following :

$$2\text{Area} = T\int V\rho d\rho = TV\alpha\beta.\int dt$$

which will be easily understood later.]

Again,

$$\rho = \alpha t + \frac{\beta}{t} \text{ or } \rho = \alpha \tan x + \beta \cot x$$

evidently represents a hyperbola referred to its asymptotes. [If t represent time, the sectorial area traced out is proportional to $\log t$, taken between proper limits.] Thus, also, the equation

$$\rho = \alpha(t + \sin t) + \beta \cos t$$

in which α and β are of equal lengths, and at right angles to one another, represents a cycloid. The origin is at the middle point of the axis (2β) of the curve. [It may be added that, if t represent *time*, this equation shows the motion of the tracing point, provided the generating circle rolls uniformly, revolving at the rate of a radian per second.]

When the lengths of α, β are not equal, this equation gives the cycloid distorted by elongation of its ordinates or abscissae : *not* a trochoid. The equation of a trochoid may be written

$$\rho = \alpha(et + \sin t) + \beta \cos t$$

e being greater or less than 1 as the curve is prolate or curtate. The lengths of α and β are still taken as equal.

But, so far as we have yet gone with the explanation of the calculus, as we are not prepared to determine the lengths or inclinations of vectors, we can investigate only a very limited class of the properties of curves, represented by such equations as those above written.

(l) We may now, in extension of the statement in §29, make the obvious remark that

$$\rho = \sum p\alpha$$

(where, as in §23, the number of vectors, α, can always be reduced to *three*, at most) is the equation of a curve in space, if the numbers p_1, p_2, &c. are functions of one indeterminate. In such a case the equation is sometimes written

$$\rho = \phi(t)$$

But, if p_1, p_2, &c. be functions of *two* indeterminates, the locus of the extremity of ρ is a *surface*; whose equation is sometimes written

$$\rho = \phi(t, u)$$

[It may not be superfluous to call the reader's attention to the fact that, in these equations, $\phi(t)$ or $\phi(t, u)$ is necessarily a vector expression, since it is equated to a vector, ρ.]

(m) Thus the equation

$$\rho = \alpha \cos t + \beta \sin t + \gamma t \tag{1}$$

belongs to a helix,

In Axiom we can draw this with the commands:

```
draw(a*cos(t)+b*sin(t)+c*u,[t=0..1,u=0..1]
```

tpdhere

while

$$\rho = \alpha \cos t + \beta \sin t + \gamma u \qquad (2)$$

represents a cylinder whose generating lines are parallel to γ,

```
draw(a*cos(t)+b*sin(t)+c*u,[t=0..1,u=0..1]
```

tpdhere

and whose base is the ellipse

$$\rho = \alpha \cos t + \beta \sin t$$

The helix above lies wholly on this cylinder.

```
draw(a*cos(t)+b*sin(t)+c*u,[t=0..1,u=0..1]
```

tpdhere

Contrast with (2) the equation

$$\rho = u(\alpha \cos t + \beta \sin t + \gamma) \qquad (3)$$

which represents a cone of the second degree

```
draw(a*cos(t)+b*sin(t)+c*u,[t=0..1,u=0..1]
```

tpdhere

made up, in fact, of all lines drawn from the origin to the ellipse

$$\rho = \alpha \cos t + \beta \sin t + \gamma$$

```
draw(a*cos(t)+b*sin(t)+c*u,[t=0..1,u=0..1]
```

tpdhere

If, however, we write

$$\rho = u(\alpha \cos t + \beta \sin t + \gamma t)$$

we form the equation of the transcendental cone whose vertex is at the origin, and on which lies the helix (1).

```
draw(a*cos(t)+b*sin(t)+c*u,[t=0..1,u=0..1]

tpdhere
```

In general

$$\rho = u\phi(t)$$

is the cone whose vertex is the origin, and on which lies the curve

$$\rho = \phi(t)$$

while

$$\rho = \phi(t) + u\alpha$$

is a cylinder, with generating lines parallel to α, standing on the same curve as base.

Again,

$$\rho = p\alpha + q\beta + r\gamma$$

with a condition of the form

$$ap^2 + bq^2 + cr^2 = 1$$

belongs to a central surface of the second order, of which α, β, γ are the directions of conjugate diameters. If a, b, c be all positive, the surface is an ellipsoid.

32. In Example (f) above we performed an operation equivalent to the differentiation of a vector with reference to a single *numerical* variable of which it was given as an explicit function. As this process is of very great use, especially in quaternion investigations connected with the motion of a particle or point; and as it will afford us an opportunity of making a preliminary step towards overcoming the novel difficulties which arise in quaternion differentiation; we will devote a few sections to a more careful, though very elementary, exposition of it.

33. It is a striking circumstance, when we consider the way in which Newton's original methods in the Differential Calculus have been decried, to find that Hamilton was *obliged* to employ them, and not the more modern forms, in order to overcome the characteristic difficulties of quaternion differentiation. Such a thing as *a differential coefficient has absolutely no meaning in quaternions*, except in those special cases in which we are dealing with degraded quaternions, such as numbers, Cartesian coordinates, &c. But a quaternion expression has always a *differential*, which is, simply, what Newton called a *fluxion*.

As with the Laws of Motion, the basis of Dynamics, so with the foundations of the Differential Calculus ; we are gradually coming to the conclusion that Newton s system is the best after all.

34. Suppose ρ to be the vector of a curve in space. Then, generally, ρ may be expressed as the sum of a number of terms, each of which is a multiple of a constant vector by a function of some *one* indeterminate; or, as in §31 (l), if P be a point on the curve,

$$\overline{OP} = \rho = \phi(t)$$

And, similarly, if Q be *any other* point on the curve,

$$\overline{OQ} = \rho_1 = \rho + \delta\rho = \phi(t_1) = \phi(t + \delta t)$$

where δt is *any number whatever.*

The vector-chord $\overline{PQ}$ is therefore, rigorously,

$$\delta p = \rho_1 - \rho = \phi(t + \delta t) - \phi t$$

35. It is obvious that, in the present case, *because the vectors involved in ϕ are constant, and their numerical multipliers alone vary*, the expression $\phi(t+\delta t)$ is, by Taylor's Theorem, equivalent to

$$\phi(t) + \frac{d\phi(t)}{dt}\delta t + \frac{d^2\phi(t)}{dt^2}\frac{(\delta t)^2}{1\ .\ 2} + \ldots$$

Hence,

$$\delta\rho = \frac{d\phi(t)}{dt}\delta t + \frac{d^2\phi(t)}{dt^2}\frac{(\delta t)^2}{1\ .\ 2} + \&\text{c.}$$

And we are thus entitled to write, when δt has been made indefinitely small,

$$\text{Limit}\left(\frac{\delta p}{\delta t}\right)_{\delta t=0} = \frac{d\rho}{dt} = \frac{d\phi(t)}{dt} = \phi^{'}(t)$$

In such a case as this, then, we are permitted to differentiate, or to form the differential coefficient of, a vector, according to the ordinary rules of the Differential Calculus. But great additional insight into the process is gained by applying Newton's method.

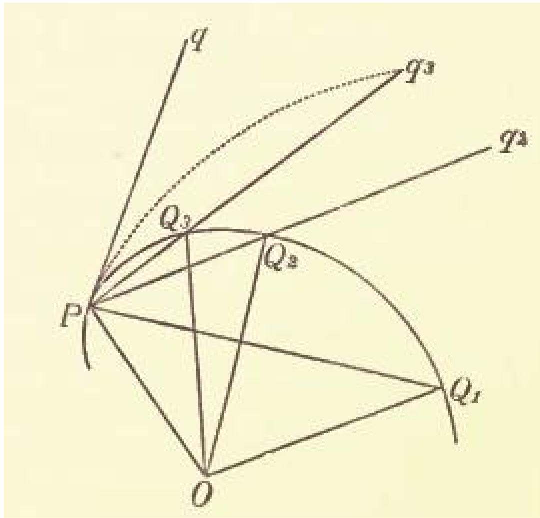
36. Let $\overline{OP}$ be

$$\rho = \phi(t)$$

and $overline OQ_1$

$$\rho_1 = \phi(t + dt)$$

where dt is any number whatever.



The number t may here be taken as representing *time*, i.e. we may suppose a point to move along the curve in such a way that the value of t for the vector of the point P of the curve denotes the interval which has elapsed (since a fixed epoch) when the moving point has reached the extremity of that vector. If, then, dt represent any interval, finite or not, we see that

$$\overline{OQ}_1 = \phi(t + dt)$$

will be the vector of the point after the additional interval dt.

But this, in general, gives us little or no information as to the velocity of the point at P. We shall get a better approximation by halving the interval dt, and finding Q_2 , where $\overline{OQ}_2 = \phi(t + \frac{1}{2}dt)$, as the position of the moving point at that time. Here the vector virtually described in $\frac{1}{2}dt$ is $\overline{PQ}_2$. To find, on this supposition, the vector described in dt, we must double $\overline{PQ}_2$, and we find, as a second approximation to the vector which the moving point would have described in time dt, if it had moved for that period in the direction and with the velocity it had at P,

$$\begin{aligned} \overline{Pq}_2 = 2\overline{PQ}_2 &= 2(\overline{OQ}_2 - \overline{OP}) \\ &= 2\{\phi(t + \tfrac{1}{2}dt) - \phi(t)\} \end{aligned}$$

The next approximation gives

$$\begin{aligned} \overline{Pq}_3 = 3\overline{PQ}_3 &= 3(\overline{OQ}_3 - \overline{OP}) \\ &= 3\{\phi(t + \tfrac{1}{3}dt) - \phi(t)\} \end{aligned}$$

And so on, each step evidently leading us nearer the sought truth. Hence, to find the vector which would have been described in time dt had the circumstances of the motion at P remained undisturbed, we must find the value of

$$d\rho = \overline{Pq} = L_{x=\infty} x \left\{ \phi\left(t + \frac{1}{x}dt\right) - \phi(t) \right\}$$

We have seen that in this particular case we may use Taylor's Theorem. We have, therefore,

$$\begin{aligned} d\rho &= L_{x=\infty} x \left\{ \phi^{'}(t)\tfrac{1}{x}dt + \phi^{''}(t)\tfrac{1}{x^2}\tfrac{(dt)^2}{1\ .\ 2} + \&\mathrm{c} \right\} \\ &= \phi^{'}(t)dt \end{aligned}$$

And, if we choose, we may now write

$$\frac{d\rho}{dt} = \phi^{'}(t)$$

37. But it is to be most particularly remarked that in the whole of this investigation no regard whatever has been paid to the magnitude of dt. The question which we have now answered may be put in the form – *A point describes a given curve in a given manner. At any point of its path its motion suddenly ceases to be accelerated. What space will it describe in a definite interval?* As Hamilton well observes, this is, for a planet or comet, the case of a 'celestial Atwood's machine'.

38. If we suppose the variable, in terms of which ρ is expressed, to be the arc, s, of the curve measured from some fixed point, we find as before

$$d\rho = \phi^{'}(x)ds$$

From the very nature of the question it is obvious that the length of dp must in this case be ds, so that $\phi^{'}(s)$ is necessarily a unit-vector. This remark is of importance, as we shall see later; and it may therefore be useful to obtain afresh the above result without any reference to time or velocity.

39. Following strictly the process of Newton s VIIth Lemma, let us describe on Pq_2 an arc similar to PQ_2, and so on. Then obviously, as the subdivision of ds is carried farther, the new arc (whose length is always ds) more and more nearly (and without limit) coincides with the line which expresses the corresponding approximation to dp.

40. As additional examples let us take some well-known *plane* curves; and first the hyperbola (§31 (k))

$$\rho = \alpha t + \frac{\beta}{t}$$

Here

$$d\rho = \left(\alpha - \frac{\beta}{t^2}\right) dt$$

This shows that the tangent is parallel to the vector

$$\alpha t - \frac{\beta}{t}$$

In words, *if the vector (from the centre) of a point in a hyperbola be one diagonal of a parallelogram, two of whose sides coincide with the asymptotes, the other* *diagonal is parallel to the tangent at the point, and cuts off a constant area from the space between the asymptotes.* (For the sides of this triangular area are t times the length of α, and $1/t$ times the length of β, respectively; the angle between them being constant.)

Next, take the cycloid, as in §31 (k),

$$\rho = \alpha(t + \sin t) + \beta \cos t$$

We have

$$d\rho = \{\alpha(1 + \cos t) - \beta \sin t\}dt$$

At the vertex

$$t = 0, \qquad \cos t = 1, \qquad \sin t = 0, \qquad \text{and } d\rho = 2\alpha dt$$

At a cusp

$$t = \pi, \qquad \cos t = -1, \qquad \sin t = 0, \qquad \text{and } d\rho = 0$$

This indicates that, at the cusp, the tracing point is (instantaneously) at rest. To find the direction of the tangent, and the form of the curve in the vicinity of the cusp, put $t = \pi + \tau$, where powers of τ above the second are omitted. We have

$$d\rho = \beta\tau dt + \frac{\alpha\tau^2}{2}dt$$

so that, at the cusp, the tangent is parallel to β. By making the same substitution in the expression for ρ, we find that the part of the curve near the cusp is a semicubical parabola,

$$\rho = \alpha(\pi + \tau^3/6) - \beta(1 - \tau^2/2)$$

or, if the origin be shifted to the cusp ($\rho = \pi\alpha - \beta$),

$$\rho = \alpha\tau^3/6 + \beta\tau^2/2$$

41. Let us reverse the first of these questions, and *seek the envelope of a line which cuts off from two fixed axes a triangle of constant area.*

If the axes be in the directions of α and β, the intercepts may evidently be written αt and $\frac{\beta}{t}$. Hence the equation of the line is (§30)

$$\rho = \alpha t + x\left(\frac{\beta}{t} - \alpha t\right)$$

The condition of envelopment is, obviously, (see Chap. IX.)

$$d\rho = 0$$

This gives $0 = \left\{\alpha - x\left(\frac{\beta}{t^2} + \alpha\right)\right\} dt + \left(\frac{\beta}{t} - \alpha t\right) dx$ †

Hence $$(1-x)dt - tdx = 0$$

and $$-\tfrac{x}{t^2}dt + \tfrac{dx}{t} = 0$$

From these, at once, $x = \frac{1}{2}$, since dx and dt are indeterminate. Thus the equation of the envelope is

$$\begin{aligned}\rho &= \alpha t + \tfrac{1}{2}\left(\tfrac{\beta}{t} - \alpha t\right)\\ &= \tfrac{1}{2}\left(\alpha t + \tfrac{\beta}{t}\right)\end{aligned}$$

the hyperbola as before; α, β being portions of its asymptotes.

42. It may assist the student to a thorough comprehension of the above process, if we put it in a slightly different form. Thus the equation of the enveloping line may be written

$$\rho = \alpha t(1-x) + \beta\frac{x}{t}$$

which gives

$$d\rho = 0 = \alpha d\{t(1-x)\} + \beta d\left(\frac{x}{t}\right)$$

Hence, as α is not parallel to β, we must have

$$d\{t(1-x)\} = 0, \qquad d\left(\frac{x}{t}\right) = 0$$

and these are, when expanded, the equations we obtained in the preceding section.

43. For farther illustration we give a solution not directly employing the differential calculus. The equations of any two of the enveloping lines are

$$\rho = \alpha t + x\left(\frac{\beta}{t} - \alpha t\right)$$

$$\rho = \alpha t_1 + x_1\left(\frac{\beta}{t_1} - \alpha t_1\right)$$

t and t_1 being given, while x and x_1 are indeterminate.

† Here we have opportunity for a remark (very simple indeed, but) of the utmost importance. *We are not to equate separately to zero the coefficients of dt and dx*; for we must remember that this equation is of the form

$$0 = p\alpha + q\beta$$

where p and q are numbers; and that, so long as α and β are actual and non-parallel vectors, the existence of such an equation requires (§24)

At the point of intersection of these lines we have (§26),

$$\left.\begin{array}{rcl} t(1-x) & = & t_1(1-x_1) \\ \frac{x}{t} & = & \frac{x_1}{t_1} \end{array}\right\}$$

These give, by eliminating x_1

$$t(1-x) = t_1\left(1 - \frac{t_1}{t}x\right)$$

or
$$x = \tfrac{t}{t_1+t}$$

Hence the vector of the point of intersection is

$$\rho = \frac{\alpha t t_1 + \beta}{t_1 + t}$$

and thus, for the ultimate intersections, where $L\frac{t_1}{t} = 1$,

$$\rho = \frac{1}{2}\left(\alpha t + \frac{\beta}{t}\right) \text{ as before}$$

COR. If. instead of the *ultimate* intersections, we consider the intersections of pairs of these lines related by some law, we obtain useful results. Thus let

$$tt_1 = 1$$

$$\rho = \frac{\alpha + \beta}{t + \frac{1}{t}}$$

or the intersection lies in the diagonal of the parallelogram on α, β.

If $t_1 = mt$, where m is constant,

$$\rho = \frac{mt\alpha + \frac{\beta}{t}}{m+1}$$

But we have also $x = \frac{1}{m+1}$

Hence *the locus of a point which divides in a given ratio a line cutting off a given area from two fixed axes, is a hyperbola of which these axes are the asymptotes.*

If we take either

$$tt_1(t + t_1) = \text{constant, or } \frac{t^2 t_1^2}{t + t_1} = \text{constant}$$

the locus is a parabola; and so on.

It will be excellent practice for the student, at this stage, to work out in detail a number of similar questions relating to the envelope of, or the locus of the intersection of selected pairs from, a series of lines drawn according to a given law. And the process may easily be extended to planes. Thus, for instance, we may form the general equation of planes which cut off constant tetrahedra from the axes of coordinates. Their envelope is a surface of the third degree whose equation may be written

$$\rho = x\alpha + y\beta + z\gamma$$

where $$xyz = \alpha^3$$

Again, find the locus of the point of intersection of three of this group of planes, such that the first intercepts on β and γ, the second on γ and α, the third on α and β, lengths all equal to one another, &c. But we must not loiter with such simple matters as these.

44. The reader who is fond of Anharmonic Ratios and Trans versals will find in the early chapters of Hamilton's *Elements of Quaternions* an admirable application of the composition of vectors to these subjects. The Theory of Geometrical Nets, in a plane, and in space, is there very fully developed; and the method is shown to include, as particular cases, the corresponding processes of Grassmann's *Ausdehnungslehre* and Möbius' *Barycentrische Calcul*. Some very curious investigations connected with curves and surfaces of the second and third degrees are also there founded upon the composition of vectors.

3.3 Examples To Chapter 1.

1. The lines which join, towards the same parts, the extremities of two equal and parallel lines are themselves equal and parallel. (*Euclid*, I. xxxiii.)

2. Find the vector of the middle point of the line which joins the middle points of the diagonals of any quadrilateral, plane or gauche, the vectors of the corners being given; and so prove that this point is the mean point of the quadrilateral.

If two opposite sides be divided proportionally, and two new quadrilaterals be formed by joining the points of division, the mean points of the three quadrilaterals lie in a straight line.

Show that the mean point may also be found by bisecting the line joining the middle points of a pair of opposite sides.

3. Verify that the property of the coefficients of three vectors whose extremities are in a line (§30) is not interfered with by altering the origin.

4. If two triangles ABC, abc, be so situated in space that Aa, Bb, Cc meet in a point, the intersections of AB, ab, of BC, bc, and of CA, ca, lie in a straight line.

5. Prove the converse of 4, i.e. if lines be drawn, one in each of two planes, from any three points in the straight line in which these planes meet, the two triangles thus formed are sections of a common pyramid.

6. If five quadrilaterals be formed by omitting in succession each of the sides of any pentagon, the lines bisecting the diagonals of these quadrilaterals meet in a point. (H. Fox Talbot.)

7. Assuming, as in §7, that the operator

$$\cos\theta + \sqrt{-1}\sin\theta$$

turns any radius of a given circle through an angle θ in the positive direction of rotation, without altering its length, deduce the ordinary formulae for $\cos(A+B)$, $\cos(A-B)$, $\sin(A+B)$, and $\sin(A-B)$, in terms of sines and cosines of A and B.

8. If two tangents be drawn to a hyperbola, the line joining the centre with their point of intersection bisects the lines join ing the points where the tangents meet the asymptotes : and the secant through the points of contact bisects the intercepts on the asymptotes.

9. Any two tangents, limited by the asymptotes, divide each other proportionally.

10. If a chord of a hyperbola be one diagonal of a parallelogram whose sides are parallel to the asymptotes, the other diagonal passes through the centre.

11. Given two points A and B, and a plane, C. Find the locus of P, such that if AP cut C in Q, and BP cut C in R, $\overline{QR}$ may be a given vector.

12. Show that $\qquad \rho = x^2\alpha + y^2\beta + (x+y)^2\gamma$
is the equation of a cone of the second degree, and that its section by the plane

$$\rho = \frac{p\alpha + q\beta + r\gamma}{p+q+r}$$

is an ellipse which touches, at their middle points, the sides of the triangle of whose corners α, β, γ are the vectors. (Hamilton, *Elements*, p. 96.)

13. The lines which divide, proportionally, the pairs of opposite sides of a gauche quadrilateral, are the generating lines of a hyperbolic paraboloid. (*Ibid.* p. 97.)

14. Show that $\qquad \rho = x^3\alpha + y^3\beta + z^3\gamma$
where $\qquad x + y + z = 0$
represents a cone of the third order, and that its section by the plane

$$\rho = \frac{p\alpha + q\beta + r\gamma}{p+q+r}$$

is a cubic curve, of which the lines

$$\rho = \frac{p\alpha + q\beta}{p+q}, \text{ \&c}$$

are the asymptotes and the three (real) tangents of inflection. Also that the mean point of the triangle formed by these lines is a conjugate point of the curve. Hence that the vector $\alpha+\beta+\gamma$ is a conjugate ray of the cone. (*Ibid.* p. 96.)

3.4 Products And Quotients of Vectors

45. We now come to the consideration of questions in which the Calculus of Quaternions differs entirely from any previous mathematical method; and here we shall get an idea of what a Quaternion is, and whence it derives its name. These questions are fundamentally involved in the novel use of the symbols of multiplication and division. And the simplest introduction to the subject seems to be the consideration of the quotient, or ratio, of two vectors.

46. If the given vectors be parallel to each other, we have already seen (§22) that either may be expressed as a numerical multiple of the other; the multiplier being simply the ratio of their lengths, taken positively if they have similar currency, negatively if they run opposite ways.

47. If they be not parallel, let $\overline{OA}$ and $\overline{OB}$ be drawn parallel and equal to them from any point O; and the question is reduced to finding the value of the ratio of two vectors drawn from the same point. Let us first find *upon how many distinct numbers this ratio depends.*

We may suppose $\overline{OA}$ to be changed into $\overline{OB}$ by the following successive processes.

1st. Increase or diminish the length of $\overline{OA}$ till it becomes equal to that of $\overline{OB}$. For this only one number is required, viz. the ratio of the lengths of the two vectors. As Hamilton remarks, this is a positive, or rather a *signless*, number.

2nd. Turn $\overline{OA}$ about O, in the common plane of the two vectors, until its direction coincides with that of $\overline{OB}$, and (remembering the effect of the first operation) we see that the two vectors now coincide or become identical. To specify this operation three numbers are required, viz. two angles (such as node and inclination in the case of a planet's orbit) to fix the plane in which the rotation takes place, and *one* angle for the amount of this rotation.

Thus it appears that the ratio of two vectors, or the multiplier required to change one vector into another, in general depends upon *four* distinct numbers, whence the name QUATERNION.

A quaternion q is thus *defined* as expressing a relation

$$\beta = q\alpha$$

between two vectors α, β. By what precedes, the vectors α, β, which serve for the definition of a given quaternion, must be in a given plane, at a given inclination to each other, and with their lengths in a given ratio ; but it is to be noticed that they may be *any* two such vectors. [*Inclination* is understood to include sense, or currency, of rotation from α to β.]

The particular case of perpendicularity of the two vectors, where their quotient is a vector perpendicular to their plane, is fully considered below; §§64, 65, 72, &c.

48. It is obvious that the operations just described may be performed, with the same result, in the opposite order, being perfectly independent of each other. Thus it appears that a quaternion, considered as the factor or agent which changes one definite vector into another, may itself be decomposed into two factors of which the order is immaterial.

The *stretching* factor, or that which performs the first operation in §47, is called the TENSOR, and is denoted by prefixing T to the quaternion considered.

The *turning factor*, or that corresponding to the second operation in §47, is called the VERSOR, and is denoted by the letter U prefixed to the quaternion.

49. Thus, if $\overline{OA} = \alpha$, $\overline{OB} = \beta$, and if q be the quaternion which changes α to β, we have

$$\beta = q\alpha$$

which we may write in the form

$$\frac{\beta}{\alpha} = q, \text{ or } \beta\alpha^{-1} = q$$

if we agree to *define* that

$$\frac{\beta}{\alpha}\alpha = \beta\alpha^{-1}\alpha = \beta$$

Here it is to be particularly noticed that we write q *before* α to signify that α is multiplied by (or operated on by) q, not q multiplied by α.

This remark is of extreme importance in quaternions, for, as we shall soon see, the Commutative Law does not generally apply to the factors of a product.

We have also, by §§47, 48,

$$q = TqUq = UqTq$$

where, as before, Tq depends merely on the relative lengths of α and β, and Uq depends solely on their directions.

Thus, if α_1 and β_1 be vectors of unit length parallel to α and β respectively,

$$T\frac{\beta_1}{\alpha_1} = T\beta_1/T\alpha_1 = 1,\ U\frac{\beta_1}{\alpha_1} = U\beta_1/U\alpha_1 = U\frac{\beta}{\alpha}$$

As will soon be shown, when α is perpendicular to β, i.e. when the versor of the quotient is quadrantal, it is a unit-vector.

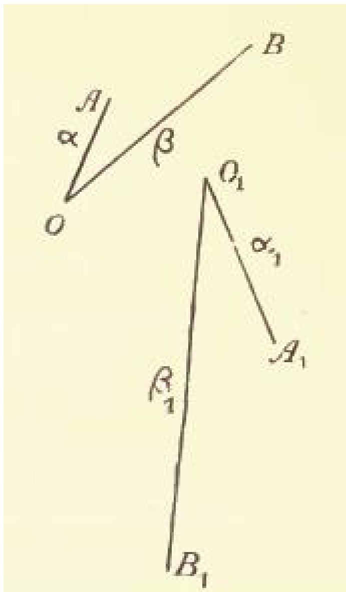
50. We must now carefully notice that the quaternion which is the quotient when β is divided by α in no way depends upon the *absolute* lengths, or directions, of these vectors. Its value will remain unchanged if we substitute for them any other pair of vectors which

(1) have their lengths in the same ratio,

(2) have their common plane the same or parallel,

and (3) make the same angle with each other.

Thus in the annexed figure



$$\frac{O_1B_1}{O_1A_1} = \frac{\overline{OB}}{\overline{OA}}$$

if, and only if,

(1) $\frac{O_1B_1}{O_1A_1} = \frac{OB}{OA}$

(2) plane AOB parallel to plane $A_1O_1B_1$

(3) $\angle AOB = \angle A_1O_1B_1$

[Equality of angles is understood to include concurrency of rotation. Thus in the annexed figure the rotation about an axis drawn upwards from the plane is negative (or clock- wise) from OA to OB, and also from O_1A_1 to O_1B_1.]

It thus appears that if

$$\beta = q\alpha,\ \delta = q\gamma$$

the vectors α, β, γ, δ are parallel to one plane, and may be repre sented (in a highly extended sense) as *proportional* to one another, thus: –

$$\beta : \alpha = \delta : \gamma$$

And it is clear from the previous part of this section that this may be written not only in the form

$$\alpha : \beta = \gamma : \delta$$

but also in either of the following forms: –

$$\gamma : \alpha = \delta : \beta$$

$$\alpha : \gamma = \beta : \delta$$

While these proportions are true as equalities of ratios, they do not usually imply equalities of products.

Thus, as the first of these was equivalent to the equation

$$\frac{\beta}{\alpha} = \frac{\delta}{\gamma} = q, \text{ or } \beta\alpha^{-1} = \delta\gamma^{-1} = q$$

the following three imply separately, (see next section)

$$\frac{\alpha}{\beta} = \frac{\gamma}{\delta} = q^{-1},\ \frac{\gamma}{\alpha} = \frac{\delta}{\beta} = r,\ \frac{\alpha}{\gamma} = \frac{\beta}{\delta} = r^{-1}$$

or, if we please,

$$\alpha\beta^{-1} = \gamma\delta^{-1} = q^{-1},\ \gamma\alpha^{-1} = \delta\beta^{-1} = r,\ \alpha\gamma^{-1} = \beta\delta^{-1} = r^{-1}$$

where r is a *new* quaternion, which has not necessarily anything (except its plane), in common with q.

But here great caution is requisite, for we are *not* entitled to conclude from these that

$$\alpha\delta = \beta\gamma, \text{ \&c.}$$

This point will be fully discussed at a later stage. Meanwhile we may merely *state* that from

$$\frac{\alpha}{\beta} = \frac{\gamma}{\delta}, \text{ or } \frac{\beta}{\alpha} = \frac{\delta}{\gamma}$$

we are entitled to deduce a number of equivalents such as

$$\alpha\beta^{-1}\delta = \gamma, \text{ or } \alpha = \gamma\delta^{-1}\beta, \text{ or } \beta^{-1}\delta = \alpha^{-1}\gamma, \text{ \&c}$$

51. The *Reciprocal* of a quaternion q is defined by the equation

$$\frac{1}{q}q = q^{-1} = 1 = q\frac{1}{q} = qqe^{-1}$$

Hence if

$$\frac{\beta}{\alpha} = q, \text{ or}$$

$$\beta = q\alpha$$

we must have

$$\frac{\alpha}{\beta} = \frac{1}{q} = q^{-1}$$

For this gives

$$\frac{\alpha}{\beta}\beta = q^{-1}q\alpha$$

and each member of the equation is evidently equal to α. Or thus: –

$$\beta = q\alpha$$

Operate *by* q^{-1}

$$q^{-1}\beta = \alpha$$

Operate *on* β^{-1}

$$q^{-1} = \alpha\beta^{-1} = \frac{\alpha}{\beta}$$

Or, we may reason thus: – since q changes $\overline{OA}$ to $\overline{OA}$, q^{-1} must change $\overline{OB}$ to $\overline{OA}$, and is therefore expressed by $\frac{\alpha}{\beta}$ (§49).

The tensor of the reciprocal of a quaternion is therefore the reciprocal of the tensor; and the versor differs merely by the *reversal* of its representative angle. The versor, it must be remembered, gives the plane and angle of the turning – it has nothing to do with the extension.

[*Remark.* In §§49–51, above, we had such expressions as $\frac{\beta}{\alpha} = \beta\alpha^{-1}$. We have also met with $\alpha^{-1}\beta$. Cayley suggests that this also may be written in the ordinary fractional form by employing the following distinctive notation: –

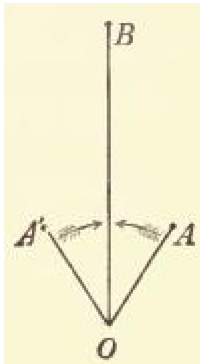
$$\frac{\beta}{\alpha} = \beta\alpha^{-1} = \frac{\beta|}{|\alpha}, \; \alpha^{-1}\beta = \frac{|\beta}{\alpha|}$$

(It might, perhaps, be even simpler to use the *solidus* as recommended by Stokes, along with an obviously correlative type:– thus,

$$\frac{\beta}{\alpha} = \beta\alpha^{-1} = \beta/\alpha, \; \alpha^{-1}\beta = \alpha\beta$$

I have found such notations occasionally convenient for private work, but I hesitate to introduce changes unless they are abso lutely required. See remarks on this point towards the end of the *Preface to the Second Edition* reprinted above.]

52. The *Conjugate* of a quaternion q, written Kq, has the same tensor, plane, and angle, only the angle is taken the reverse way; or the versor of the conjugate is the reciprocal of the versor of the quaternion, or (what comes to the same thing) the versor of the reciprocal.



Thus, if OA, OB, OA' , lie in one plane, and if $OA' = OA$, and $\angle A'OB = \angle BOA$, we have

$$\frac{\overline{OB}}{\overline{OA}} = q$$

, and

$$\frac{\overline{OB}}{\overline{OA'}} = \text{ congugate of } q = Kq$$

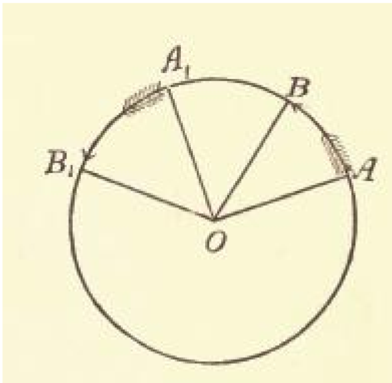
By last section we see that

$$Kq = (Tq)^2 q^{-1}$$

Hence $$qKq = Kqq = (Tq)^2$$

This proposition is obvious, if we recollect that the tensors of q and Kq are equal, and that the versors are such that either *annuls* the effect of the other; while the order of their application is indifferent. The joint effect of these factors is therefore merely to multiply twice over by the common tensor.

53. It is evident from the results of §50 that, if α and β be of equal length, they may be treated as of unit-length so far as their quaternion quotient is concerned. This quotient is therefore a versor (the tensor being unity) and may be represented indifferently by any one of an infinite number of concurrent arcs of given length lying on the circumference of a circle, of which the two vectors are radii. This is of considerable importance in the proofs which follow.



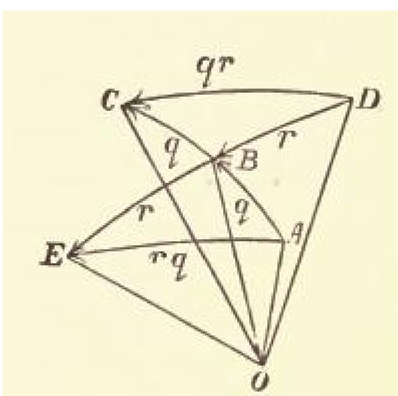
Thus the versor $\frac{\overline{OB}}{\overline{OA}}$ may be represented in magnitude, plane, and currency of rotation (§50) by the arc AB, which may in this extended sense be written $\widehat{AB}$.

And, similarly, the versor $\frac{\overline{OB_1}}{\overline{OA_1}}$ may be represented by $\widehat{A_1B_1}$ which is equal to (and concurrent with) $\widehat{AB}$ if

$$\angle A_1OB_1 = \angle AOB$$

i.e. if the versors are *equal*, in the quaternion meaning of the word.

54. By the aid of this process, when a versor is represented as an arc of a great circle on the unit-sphere, we can easily prove that *quaternion multiplication is not generally commutative*.



Thus let q be the versor $\widehat{AB}$ or $\frac{\overline{OB}}{\overline{OA}}$, where O is the centre of the sphere.

Take $\widehat{BC} = \widehat{AB}$, (which, it must be remembered, makes the points A, B, C, lie in one great circle), then q may also be represented by $\frac{\overline{OC}}{\overline{OB}}$.

In the same way any other versor r may be represented by $\overset{\frown}{DB}$ or $\overset{\frown}{BE}$ and by $\frac{\overline{OB}}{\overline{OD}}$ or $\frac{\overline{OE}}{\overline{OB}}$.

[The line OB in the figure is definite, and is given by the intersection of the planes of the two versors.]

Now $r\overline{OD} = \overline{OB}$, and $q\overline{OB} = \overline{OC}$.

Hence $qr\overline{OD} = \overline{OC}$,

or $qr = \frac{\overline{OC}}{\overline{OD}}$, and may therefore be represented by the arc $\overset{\frown}{DC}$ of a great circle.

But rq is easily seen to be represented by the arc $\overset{\frown}{AE}$.

For $q\overline{OA} = \overline{OB}$, and $r\overline{OB} = \overline{OE}$,

whence $rq\overline{OA} = \overline{OE}$. and $rq = \frac{\overline{OE}}{\overline{OA}}$.

Thus the versors rq and qr, though represented by arcs of equal length, are not generally in the same plane and are therefore unequal: unless the planes of q and r coincide.

Remark. We see that we have assumed, or defined, in the above proof, that $q.r\alpha = qr.\alpha$. and $r.q\alpha = rq.\alpha$ in the special case when $q\alpha$, $r\alpha$, $q.r\alpha$ and $r.q\alpha$ are all *vectors*.

55. Obviously $\overset{\frown}{CB}$ is Kq, $\overset{\frown}{BD}$ is Kr, and $\overset{\frown}{CD}$ is $K(qr)$. But $\overset{\frown}{CD} = \overset{\frown}{BD}.\overset{\frown}{CB}$ as we see by applying both to OC. This gives us the very important theorem

$$K(qr) = Kr.Kq$$

i.e. *the conjugate of the product of two versors is the product of their conjugates in inverted order*. This will, of course, be extended to any number of factors as soon as we have proved the associative property of multiplication. (§58 below.)

56. The propositions just proved are, of course, true of quater nions as well as of versors; for the former involve only an additional numerical factor which has reference to the length merely, and not the direction, of a vector (§48), and is therefore commutative with all other factors.

57. Seeing thus that the commutative law does not in general hold in the multiplication of quaternions, let us enquire whether the Associative Law holds generally. That is if p, q, r be three quaternions, have we

$$p.qr = pq.r?$$

This is, of course, obviously true if p, q, r be numerical quantities, or even any of the imaginaries of algebra. But it cannot be con sidered as a truism for symbols which do not in general give

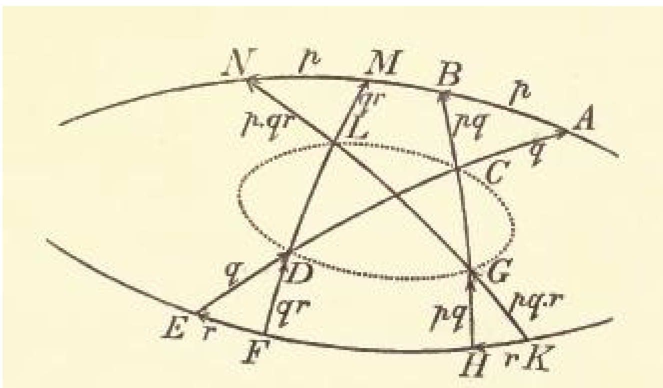
$$pq = qp$$

We have assumed it, in definition, for the special case when r, qr, and pqr are all vectors. (§54.) But we are not entitled to assume any more than is absolutely required to make our definitions complete.

58. In the first place we remark that p, q, and r may be considered as versors only, and therefore represented by arcs of great circles on the unit sphere, for their tensors may obviously (§48) be divided out from both sides, being commutative with the versors.

Let $\frown{AB} = p$, $\frown{ED} = \frown{CA} = q$, and $\frown{FE} = r$.

Join BC and produce the great circle till it meets EF in H, and make $\frown{KH} = \frown{FE} = r$, and $\frown{HG} = \frown{CB} = pq$ (§54).



Join GK. Then $\frown{KG} = \frown{HG}.\frown{KH} = pq.r$.

Join FD and produce it to meet AB in M. Make

$$\frown{LM} = \frown{FD}, \text{ and } \frown{MN} = \frown{AB}$$

and join NL. Then

$$\frown{LN} = \frown{MN}.\frown{LM} = p.qr$$

.

Hence to show that $p.qr = pq.r$

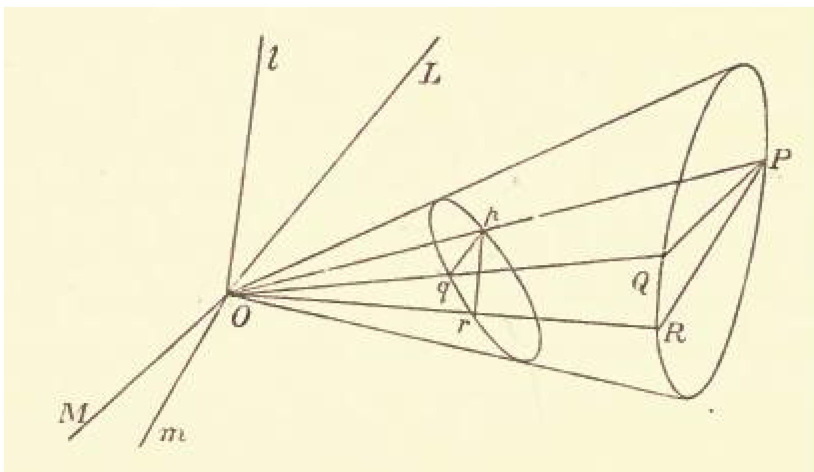
all that is requisite is to prove that LN, and KG, described as above, are *equal arcs of the same great circle*, since, by the figure, they have evidently similar currency. This is perhaps most easily effected by the help of the fundamental properties of the curves known as *Spherical Conics*. As they are not usually familiar to students, we make a slight digression for the purpose of proving these fundamental properties ; after Chasles, by whom and Magnus they were discovered. An independent proof of the associative principle will presently be indicated, and in Chapter VIII. we shall employ quaternions to give an independent proof of the theorems now to be established.

59.* DEF. *A spherical conic is the curve of intersection of a cone of the second degree with a sphere, the vertex of the cone being the centre of the sphere.*

LEMMA. If a cone have one series of circular sections, it has another series, and any two circles belonging to different series lie on a sphere. This is easily proved as follows.

Describe a sphere, A, cutting the cone in one circular section, C, and in any other point whatever, and let the side OpP of the cone meet A in p, P ; P being a point in C. Then $PO.Op$ is constant, and, therefore, since P lies in a plane, p lies on a sphere, a, passing through 0. Hence the locus, c, of p is a circle, being the intersection of the two spheres A and a.

Let OqQ be any other side of the cone, q and Q being points in c, C respectively. Then the quadrilateral $qQPp$ is inscribed in a circle (that in which its plane cuts the sphere A) and the exterior



angle at p is equal to the interior angle at Q. If OL, OM be the lines in which the plane POQ cuts the *cyclic planes* (planes through O parallel to the two series of circular sections) they are obviously parallel to pq, QP, respectively; and therefore

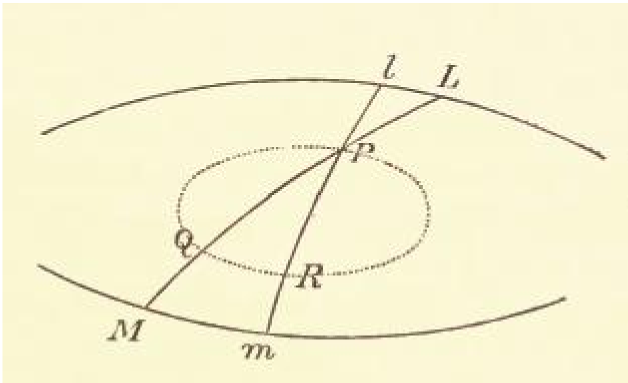
$$\angle LOp = \angle Opq = \angle OQP = \angle MOQ$$

Let any third side, OrR, of the cone be drawn, and let the plane OPR cut the cyclic planes in $0l$, Om respectively. Then, evidently,

$$\angle lOL = \angle qpr$$

$$\angle MOm = \angle QPR$$

and these angles are independent of the position of the points p and P, if Q and R be fixed points.



In the annexed section of the above space-diagram by a sphere whose centre is O, lL, Mm are the great circles which represent the cyclic planes, PQR is the spherical conic which represents the cone. The point P represents the line OpP, and so with the others. The propositions above may now be stated thus,

$$\text{Arc } PL = \text{arc } MQ$$

and, if Q and R be fixed, Mm and lL are constant arcs whatever be the position of P.

60. The application to §58 is now obvious. In the figure of that article we have

$$\overset{\frown}{FE} = \overset{\frown}{KH},\ \overset{\frown}{ED} = \overset{\frown}{CA},\ \overset{\frown}{HG} = \overset{\frown}{CB},\ \overset{\frown}{LM} = \overset{\frown}{FD}$$

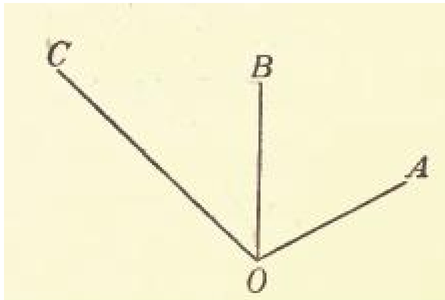
Hence L, C, G, D are points of a spherical conic whose cyclic planes are those of AB, FE. Hence also KG passes through L, and with LM intercepts on AB an arc equal to $\overset{\frown}{AB}$. That is, it passes through N, or KG and LN are arcs of the same great circle : and they are equal, for G and L are points in the spherical conic.

Also, the associative principle holds for any number of quaternion factors. For, obviously,

$$qr.st = qrs.t = \&c., \&c.,$$

since we may consider qr as a single quaternion, and the above proof applies directly.

61. That quaternion addition, and therefore also subtraction, is commutative, it is easy to show.



For if the planes of two quaternions, q and r, intersect in the line OA, we may take any vector $\overline{OA}$ in that line, and at once find two others, $\overline{OB}$ and $\overline{OC}$, such that

$$\overline{OB} = q\overline{OA}$$

and $$\overline{CO} = r\overline{OA}$$

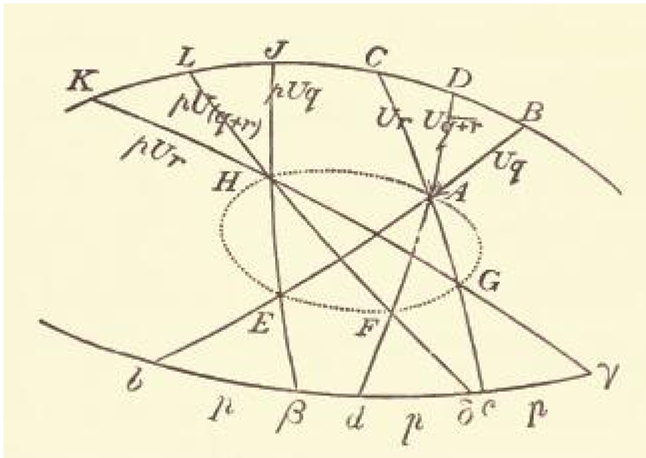
And $$(q + r)\overline{OAOB} + \overline{OC} = \overline{OC} + \overline{OB} = (r + q)\overline{OA}$$

since vector addition is commutative (§27).

Here it is obvious that $(q + r)\overline{OA}$, being the diagonal of the parallelogram on $\overline{OB}$, $\overline{OC}$, divides the angle between OB and OC in a ratio depending solely on the ratio of the lengths of these lines, i.e. on the ratio of the tensors of q and r. This will be useful to us in the proof of the distributive law, to which we proceed.

62. Quaternion multiplication, and therefore division, is distributive. One simple proof of this depends on the possibility, shortly to be proved, of representing *any* quaternion as a linear function of three given rectangular unit- vectors. And when the proposition is thus established, the associative principle may readily be deduced from it.

[But Hamilton seems not to have noticed that we may employ for its proof the properties of Spherical Conies already employed



in demonstrating the truth of the associative principle. "For continuity we give an outline of the proof by this process.

Let $\overset{\frown}{BA}$, $\overset{\frown}{CA}$ represent the versors of q and r, and be the great circle whose plane is that of p.

Then, if we take as operand the vector $\overline{OA}$, it is obvious that $U(q+r)$ will be represented by some such arc as $\overset{\frown}{DA}$ where B, D, C are in one great circle; for $(q+r)\overline{OA}$ is in the same plane as $q\overline{OA}$ and $r\overline{OA}$, and the relative magnitude of the arcs BD and DC depends solely on the tensors of q and r. Produce BA, DA, CA to meet be in b, d, c respectively, and make

$$\overset{\frown}{Eb} = \overset{\frown}{BA}, \overset{\frown}{Fd} = \overset{\frown}{DA}, \overset{\frown}{Gc} = \overset{\frown}{CA}$$

Also make $\overset{\frown}{b\beta} = \overset{\frown}{d\delta} = \overset{\frown}{c\gamma} = p$. Then E, F, G, A lie on a spherical conic of which BC and bc are the cyclic arcs. And, because $\overset{\frown}{b\beta} = \overset{\frown}{d\delta} = \overset{\frown}{c\gamma}$, $\overset{\frown}{\beta E}$, $\overset{\frown}{\delta F}$, $\overset{\frown}{\gamma G}$, when produced, meet in a point H which is also on the spherical conic (§59*). Let these arcs meet BC in J, L, K respectively. Then we have

$$\overset{\frown}{JH} = \overset{\frown}{E\beta} = pUq$$

$$\overset{\frown}{LH} = \overset{\frown}{F\delta} = pU(q+r)$$

$$\overset{\frown}{KH} = \overset{\frown}{G\gamma} = pUr$$

Also $$\overset{\frown}{LJ} = \overset{\frown}{DB}$$

and $$\overset{\frown}{KL} = \overset{\frown}{CD}$$

And, on comparing the portions of the figure bounded respectively by HKJ and by ACB we see that (when considered with reference to their effects as factors multiplying $\overline{OH}$ and $\overline{OA}$ respectively)

$pU(q4+r)$ bears the same relation to pUq and pUr
that $U(q+r)$ bears to Uq and Ur.
But $T(q+r)U(q+r) = q+r = TqUq + TrUr$.
Hence $T(q+r).pU(q+r) = Tq.pUq + Tr.pUr$;
or, since the tensors are mere numbers and commutative with all other factors,

$$p(q+r) = pq + pr$$

In a similar manner it may be proved that

$$(q+)p = qp + rp$$

And then it follows at once that

$$(p+q)(r+s) = pr + ps + qr + qs$$

where, by §61, the order of the partial products is immaterial.]

63. By similar processes to those of §53 we see that versors, and therefore also quaternions, are subject to the index-law

$$q^m.q^n = q^{m+n}$$

at least so long as m and n are positive integers.

The extension of this property to negative and fractional exponents must be deferred until we have defined a negative or fractional power of a quaternion.

64. We now proceed to the special case of *quadrantal* versors, from whose properties it is easy to deduce all the foregoing results of this chapter. It was, in fact, these properties whose invention by Hamilton in 1843 led almost intuitively to the establishment of the Quaternion Calculus. We shall content ourselves at present with an assumption, which will be shown to lead to consistent results ; but at the end of the chapter we shall show that no other assumption is possible, following for this purpose a very curious quasi-metaphysical speculation of Hamilton.

65. Suppose we have a system of three mutually perpendicular unit-vectors, drawn from one point, which we may call for shortness **i**, **j**, **k**. Suppose also that these are so situated that a positive (i.e. *left-handed*) rotation through a right angle about **i** as an axis brings **j** to coincide with **k**. Then it is obvious that positive quadrantal rotation about **j** will make **k** coincide with **i**; and, about **k**, will make **i** coincide with **j**.

For defniteness we may suppose **i** to be drawn *eastwards*, **j** *northwards*, and **k** *upwards*. Then it is obvious that a positive (left-handed) rotation about the eastward line (**i**) brings the northward line (**j**) into a vertically upward position (**k**) ; and so of the others.

66. Now the operator which turns **j** into **k** is a quadrantal versor (§53) ; and, as its axis is the vector **i**, we may call it *i*.

Thus
$$\frac{\mathbf{k}}{\mathbf{j}} = i, \text{ or } \mathbf{k} = i\mathbf{j} \tag{1}$$
Similary we may put
$$\frac{\mathbf{i}}{\mathbf{k}} = j, \text{ or } \mathbf{i} = j\mathbf{k} \tag{2}$$
and
$$\frac{\mathbf{j}}{\mathbf{i}} = k, \text{ or } \mathbf{j} = k\mathbf{i} \tag{3}$$

[It may be here noticed, merely to show the symmetry of the system we arc explaining, that if the three mutually perpendicular vectors **i**, **j**, **k** be made to revolve about a line equally inclined to all, so that **i** is brought to coincide with **j**, **j** will then coincide with **k**, and **k** with **i**: and the above equations will still hold good, only (1) will become (2), (2) will become (3), and (3) will become (1).]

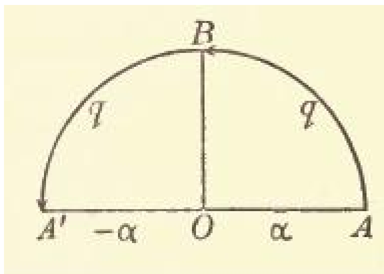
67. By the results of §50 we see that
$$\frac{-\mathbf{j}}{\mathbf{k}} = \frac{\mathbf{k}}{\mathbf{j}}$$
i.e. a southward unit- vector bears the same ratio to an upward unit-vector that the latter does to a northward one; and therefore we have

Thus
$$\frac{-\mathbf{j}}{\mathbf{k}} = i, \text{ or } -\mathbf{j} = i\mathbf{k} \tag{4}$$
Similary t
$$\frac{-\mathbf{k}}{\mathbf{i}} = j, \text{ or } -\mathbf{k} = j\mathbf{i} \tag{5}$$
and
$$\frac{-\mathbf{i}}{\mathbf{j}} = k, \text{ or } -\mathbf{i} = k\mathbf{j} \tag{6}$$

68. By (4) and (1) we have
$$-j = ik = i(ij) \text{ (by the assumption in §54) } = i^2 j$$
Hence
$$i^2 = -1 \tag{7}$$
Arid in the same way, (5) and (2) give
$$j^2 = -1 \tag{8}$$
and (6) and (3)
$$k^2 = -1 \tag{9}$$

Thus, as the directions of **i**, **j**, **k** are perfectly arbitrary, we see that *the square of every quadrantal versor is negative unity*.

[Though the following proof is in principle exactly the same as the foregoing, it may perhaps be of use to the student, in showing him precisely the nature as well as the simplicity of the step we have taken.



Let ABA' be a semicircle, whose centre is 0, and let OB be perpendicular to AOA'.

Then $\frac{\overline{OB}}{\overline{OA'}} = q$ suppose, is a quadrantal versor, and is evidently equal to $\frac{\overline{OA'}}{\overline{OB}}$; §§50, 53. Hence

$$q^2 = \frac{\overline{OA'}}{\overline{OB}} . \frac{\overline{OB}}{\overline{OA}} = \frac{\overline{OA'}}{\overline{OA}} = -1]$$

69. Having thus found that the squares of i, j, k are each equal to negative unity ; it only remains that we find the values of their products two and two. For, as we shall see, the result is such as to show that the value of any other combination whatever of i,j, k (as factors of a product) may be deduced from the values of these squares and products.

Now it is obvious that

$$\frac{\mathbf{k}}{-\mathbf{i}} = \frac{\mathbf{i}}{\mathbf{k}} = j$$

(i.e. the versor which turns a westward unit-vector into an upward one will turn the upward into an eastward unit) ; or

$$\mathbf{k} = j(-\mathbf{i}) = -j\mathbf{i} \tag{10}$$

Now let us operate on the two equal vectors in (10) by the same versor, i, and we have

$$i\mathbf{k} = i(-j\mathbf{i}) = -ij\mathbf{i}$$

But by (4) and (3)

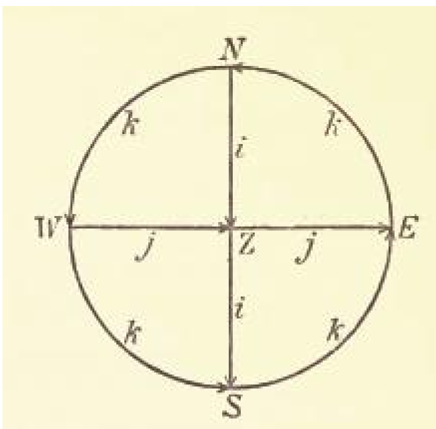
$$i\mathbf{k} = -\mathbf{j} = -k\mathbf{i}$$

Comparing these equations, we have

$$-ij\mathbf{i} = -k\mathbf{i}$$

$$\left.\begin{array}{ll}\text{or, §54 (end),} & ij = k \\ \text{and symmetry gives} & jk = i \\ & ki = j\end{array}\right\} \qquad (11)$$

The meaning of these important equations is very simple ; and is, in fact, obvious from our construction in §54 for the multiplication of versors ; as we see by the annexed figure, where we must remember that i, j, k are quadrantal versors whose planes are at right angles, so that the figure represents a hemisphere divided into quadrantal triangles. [The arrow-heads indicate the direction of each vector arc.]



Thus, to show that $ij = k$, we have, O being the centre of the sphere, N, E, S, W the north, east, south, and west, and Z the zenith (as in §65) ;

$$j\overline{OW} = \overline{OZ}$$

whence $$ij\overline{OW} = i\overline{OZ} = \overline{OS} = k\overline{OW}$$

* The negative sign, being a mere numerical factor, is evidently commutative with j indeed we may, if necessary, easily assure ourselves of the fact that to turn the negative (or reverse) of a vector through a right (or indeed any) angle, is the same thing as to turn the vector through that angle and then reverse it.

70. But, by the same figure,

$$i\overline{ON} = \overline{OZ}$$

whence $$ji\overline{ON} = j\overline{OZ} = \overline{OE} = -\overline{OW} = -k\overline{ON}.$$

71. From this it appears that

$$\left.\begin{array}{l} ji = -k \\ kj = -i \\ ik = -j \end{array}\right\} \qquad (12)$$

and thus, by comparing (11),

$$\left.\begin{array}{l} ij = -ji = k \\ jk = -kj = i \\ ki = -ik = j \end{array}\right\} \qquad (11),(12)$$

These equations, along with

$$i^2 = j^2 = k^2 = -1 \qquad ((7),(8),(9))$$

contain essentially the whole of Quaternions. But it is easy to see that, for the first group, we may substitute the single equation

$$ijk = -1 \qquad (13)$$

since from it, by the help of the values of the squares of i, j, k, all the other expressions may be deduced. We may consider it proved in this way, or deduce it afresh from the figure above, thus

$$k\overline{ON} = \overline{OW}$$

$$jk\overline{ON} = j\overline{OW} = \overline{OZ}$$

$$ijk\overline{ON} = ij\overline{OW} = i\overline{OZ} = \overline{OS} = -\overline{ON}$$

72. One most important step remains to be made, to wit the assumption referred to in §64. We have treated i, j, k simply as quadrantal versors ; and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ as unit-vectors at right angles to each other, and coinciding with the axes of rotation of these versors. But if we collate and compare the equations just proved we have

$$\left\{\begin{array}{ll} i^2 = -1 & (7) \\ \mathbf{i}^2 = -1 & (\S 9) \end{array}\right.$$

$$\left\{\begin{array}{ll} ij = k & (11) \\ i\mathbf{j} = \mathbf{k} & (1) \end{array}\right.$$

$$\left\{\begin{array}{ll} ji = -k & (11) \\ j\mathbf{i} = -\mathbf{k} & (1) \end{array}\right.$$

with the other similar groups symmetrically derived from them.

Now the meanings we have assigned to i, j, k are quite independent of, and not inconsistent with, those assigned to $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. And it is superfluous to use two sets of characters when one will suffice. Hence it appears that i, j, k may be substituted for $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$; in other words, *a unit-vector when employed as a factor may be considered as a quadrantal versor whose plane is perpendicular to the vector*. (Of course it follows that every vector can be treated as the product of a number and a quadrantal versor.) This is one of the main elements of the singular simplicity of the quaternion calculus.

73. Thus *the product, and therefore the quotient, of two perpendicular vectors is a third vector perpendicular to both*.

Hence the reciprocal (§51) of a vector is a vector which has the *opposite* direction to that of the vector, arid its length is the reciprocal of the length of the vector.

The conjugate (§52) of a vector is simply the vector reversed.

Hence, by §52, if α be a vector

$$(Ta)^2 = \alpha K\alpha = \alpha(-\alpha) = -\alpha^2$$

74. We may now see that *every versor may be represented by a power of a unit-vector.*

For, if α be any vector perpendicular to i (which is *any* definite unit-vector), $i\alpha = \beta$ is a vector equal in length to α, but perpendicular to both i and α

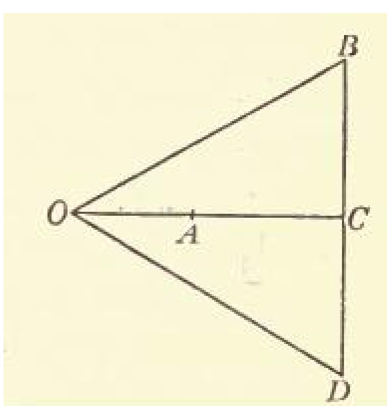
$$\begin{aligned} i^2\alpha &= -\alpha \\ \imath^3\alpha &= -i\alpha = -\beta \\ \imath^4\alpha &= -i\beta = -i^2\alpha = \alpha \end{aligned}$$

Thus, by successive applications of i, $\alpha.$ is turned round i as an axis through successive right angles. Hence it is natural to *define i^m as a versor which turns any vector perpendicular to i through m right angles in the positive direction of rotation about i as an axis.* Here m may have any real value whatever, whole or fractional, for it is easily seen that analogy leads us to interpret a negative value of m as corresponding to rotation in the negative direction.

75. From this again it follows that *any quaternion may be expressed as a power of a vector.* For the tensor and versor elements of the vector may be so chosen that, when raised to the same power, the one may be the tensor and the other the versor of the given quaternion. The vector must be, of course, perpen dicular to the plane of the quaternion.

76. And we now see, as an immediate result of the last two sections, that the index-law holds with regard to powers of a quaternion (§63).

77. So far as we have yet considered it, a quaternion has been regarded as the *product* of a tensor and a versor: we are now to consider it as a *sum.* The easiest method of so analysing it seems to be the following.



Let $\dfrac{\overline{OB}}{\overline{OA}}$ represent any quaternion. Draw BC perpendicular to OA, produced if necessary.

Then, §19, $\overline{OB} = \overline{OC} + \overline{CB}$

But, §22, $\overline{OC} = x\overline{OA}$
where x is a number, whose sign is the same as that of the cosine of $\angle AOB$.

Also, §73, since CB is perpendicular to OA,

$$\overline{CB} = \gamma\overline{OA}$$

where γ is a vector perpendicular to OA and CB, i.e. to the plane of the quaternion; and, as the figure is drawn, directed *towards* the reader.

Hence

$$\frac{\overline{OB}}{\overline{OA}} = \frac{x\overline{OA} + \gamma\overline{OA}}{\overline{OA}} = x + \gamma$$

Thus a quaternion, in general, may be decomposed into the sum of two parts, one numerical, the other a vector. Hamilton calls them the SCALAR, and the VECTOR, and denotes them respectively by the letters S and V prefixed to the expression for the quaternion.

78. Hence $q = Sq + Vq$, and if in the above example

$$\frac{\overline{OB}}{\overline{OA}} = q$$

then

$$\overline{OB} = \overline{OC} + \overline{CB} = Sq.\overline{OA} + Vq.\overline{OA}^{\dagger}$$

The equation above gives

$$\overline{OC} = Sq.\overline{OA}$$
$$\overline{CB} = Vq.\overline{OA}$$

79. If, in the last figure, we produce BC to D, so as to double its length, and join OD, we have, by §52,

$$\frac{\overline{OD}}{\overline{OA}} = Kq = SKq + VKq$$

so that $\overline{OD} = \overline{OC} + \overline{CD} = SKq.\overline{OA} + VKq.\overline{OA}$
Hence $\overline{OC} = SKq.\overline{OA}$
and $\overline{CD} = VKq.\overline{OA}$
Comparing this value of $\overline{OC}$ with that in last section, we find

$$SKq = Sq \qquad (1)$$

† The points are inserted to show that S and V apply only to q, and not to $q\overline{OA}$.

or *the scalar of the conjugate of a quaternion is equal to the scalar of the quaternion.*

Again, $\overline{CD} = -\overline{CB}$ by the figure, and the substitution of their values gives

$$VKq = -Vq \tag{2}$$

or *the vector of the conjugate of a quaternion is the vector of the quaternion reversed.*

We may remark that the results of this section are simple con sequences of the fact that the symbols S, V, K are commutative †.

Thus $SKq = KSq = Sq$,
since the conjugate of a number is the number itself; and

$$VKq = KVq = -Vq(\S 73)$$

Again, it is obvious that,

$$\sum Sq = S\sum q, \quad \sum Vq = V\sum q$$

and thence $\sum Kq = K\sum q$

80. Since any vector whatever may be represented by

$$xi + yj + zk$$

where x, y, z are numbers (or Scalars), and i, j, k may be any three non-coplanar vectors, §§23, 25 though they are usually understood as representing a rectangular system of unit-vectors and since any scalar may be denoted by w; we may write, for any quaternion q, the expression

$$q = w + xi + yj + zk(\S 78)$$

Here we have the essential dependence on four distinct numbers, from which the quaternion derives its name, exhibited in the most simple form.

And now we see at once that an equation such as

$$q^{'} = q$$

where $q^{'} = w^{'} + x^{'}i + y^{'}j + z^{'}k$
involves, of course, the *four* equations

$$w^{'} = w, x^{'} = x, y^{'} = y, z^{'} = z$$

† It is curious to compare the properties of these quaternion symbols with those of the Elective Symbols of Logic, as given in BOOLE'S wonderful treatise on the *Laws of Thought*; and to think that the same grand science of mathematical analysis, by processes remarkably similar to each other, reveals to us truths in the science of position far beyond the powers of the geometer, and truths of deductive reasoning to which unaided thought could never have led the logician.

81. We proceed to indicate another mode of proof of the distributive law of multiplication.

We have already defined, or assumed (§61), that

$$\frac{\beta}{\alpha}+\frac{\gamma}{\alpha}=\frac{\beta+\gamma}{\alpha}$$

or $$\beta\alpha^{-1}+\gamma\alpha^{-1}=(\beta+\gamma)\alpha^{-1}$$

and have thus been able to understand what is meant by adding two quaternions.

But, writing α for α^{-1}, we see that this involves the equality

$$(\beta+\gamma)\alpha=\beta\alpha+\gamma\alpha$$

from which, by taking the conjugates of both sides, we derive

$$\alpha^{'}(\beta^{'}+\gamma^{'})=\alpha^{'}\beta^{'}+\alpha^{'}\gamma^{'}(\S 55)$$

And a combination of these results (putting $\beta+\gamma$ for $\alpha^{'}$ in the latter, for instance) gives

$$\begin{aligned}(\beta+\gamma)(\beta^{'}+\gamma^{'}) &= (\beta+\gamma)\beta^{'}+(\beta+\gamma)\gamma^{'}\\ &= \beta\beta^{'}+\gamma\beta^{'}+\beta\gamma^{'}+\gamma\gamma^{'}\end{aligned}$$

by the former.

Hence the *distributive principle is true in the multiplication of vectors.*

It only remains to show that it is true as to the scalar and vector parts of a quaternion, and then we shall easily attain the general proof.

Now, if a be any scalar, α any vector, and q any quaternion,

$$(a+\alpha)q=aq+\alpha q$$

For, if β be the vector in which the plane of q is intersected by a plane perpendicular to α, we can find other two vectors, γ and δ one in each of these planes such that

$$\alpha=\frac{\gamma}{\beta},\qquad q=\frac{\beta}{\delta}$$

And, of course, a may be written $\frac{a\beta}{\beta}$; so that

$$\begin{aligned}(a+\alpha)q &= \tfrac{a\beta+\gamma}{\beta}.\tfrac{\beta}{\delta}=\tfrac{a\beta+\gamma}{\delta}\\ &= a\tfrac{\beta}{\delta}+\tfrac{\gamma}{\delta}=a\tfrac{\beta}{\delta}+\tfrac{\gamma}{\beta}.\tfrac{\beta}{\delta}\\ &= aq+\alpha q\end{aligned}$$

And the conjugate may be written

$$q^{'}(a^{'}+\alpha^{'})=q^{'}a^{'}+q^{'}\alpha^{'}(\S 55)$$

Hence, generally,

$$(a+\alpha)(b+\beta) = ab + a\beta + b\alpha + \alpha\beta$$

or, breaking up a and b each into the sum of two scalars, and α, β each into the sum of two vectors,

$$(a_1 + a_2 + \alpha_1 + \alpha_2)(b_1 + b_2 + \beta_1 + \beta_2)$$

$$= (a_1+a_2)(b_1+b_2) + (a_1+a_2)(\beta_1+\beta_2) + (b_1+b_2)(\alpha_1+\alpha_2) + (\alpha_1+\alpha_2)(\beta_1+\beta_2)$$

(by what precedes, all the factors on the right are distributive, so that we may easily put it in the form)

$$= (a_1+\alpha_1)(b_1+\beta_1) + (a_1+\alpha_1)(b_2+\beta_2) + (a_2+\alpha_2)(b_1+\beta_1) + (a_2+\alpha_2)(b_2+\beta_2)$$

Putting $a_1 + \alpha_1 = p, \quad a_2 + \alpha_2 = q, \quad b_1 + \beta_1 = r, \quad b_2 + \beta_2 = s,$
we have $(p+q)(r+s) = pr + ps + qr + qs$

82. Cayley suggests that the laws of quaternion multiplication may be derived more directly from those of vector multiplication, supposed to be already established. Thus, let α be the unit vector perpendicular to the vector parts of q and of $q^{'}$. Then let

$$\rho = q.\alpha, \quad \sigma = -\alpha.q^{'}$$

as is evidently permissible, and we have

$$p\alpha = q.\alpha\alpha = -q; \quad \alpha\sigma = -\alpha\alpha.q^{'} = q^{'}$$

so that
$$-q.q^{'} = \rho\alpha.\alpha\sigma = -\rho.\sigma$$

The student may easily extend this process.

For variety, we shall now for a time forsake the geometrical mode of proof we have hitherto adopted, and deduce some of our next steps from the analytical expression for a quaternion given in §80, and the properties of a rectangular system of unit-vectors as in §71.

We will commence by proving the result of §77 anew.

83. Let

$$\alpha = xi + yj + zk$$
$$\beta = x^{'}i + y^{'}j + z^{'}k$$

Then, because by §71 every product or quotient of i, j, k is reducible to one of them or to a number, we are entitled to assume

$$q = \frac{\beta}{\alpha} = \omega + \xi i + \eta j + \zeta k$$

where ω, ξ, η, ζ are numbers. This is the proposition of §80.

[Of course, with this expression for a quaternion, there is no necessity for a formal proof of such equations as

$$p + (q+r) = (p+q) + r$$

where the various sums are to be interpreted as in §61.

All such things become obvious in view of the properties of i, j ,k.]

84. But it may be interesting to find ω, ξ, η, ζ in terms of x, y, z, x', y' , z' . We have

$$\beta = q\alpha$$

or

$$x'i + y'j + z'k = (\omega + \xi i + \eta j + \zeta k)(xi + yj + zk)$$

$$= -(\xi x + \eta y + \zeta z) + (\omega x + \eta z - \zeta y)i + (\omega y + \zeta x - \xi z)j + (\omega z + \xi y - \eta x)k$$

as we easily see by the expressions for the powers and products of i, j, k given in §71. But the student must pay particular attention to the *order* of the factors, else he is certain to make mistakes.

This (§80) resolves itself into the four equations

$$\begin{array}{lllllllll}
0 & = & & & \xi x & + & \eta y & + & \zeta z \\
x' & = & \omega x & & & + & \eta z & - & \zeta y \\
y' & = & \omega y & - & \xi z & & & + & \zeta x \\
z' & = & \omega z & + & \xi y & - & \eta x & &
\end{array}$$

The three last equations give

$$xx' + yy' + zz' = \omega(x^2 + y^2 + z^2)$$

which determines ω.

Also we have, from the same three, by the help of the first,

$$\xi x' + \eta y' + \zeta z' = 0$$

which, combined with the first, gives

$$\frac{\xi}{yz' - zy'} = \frac{\eta}{zx' - xz'} = \frac{\zeta}{xy' - yx'}$$

and the common value of these three fractions is then easily seen to be

$$\frac{1}{x^2 + y^2 + z^2}$$

It is easy enough to interpret these expressions by means of ordinary coordinate geometry : but a much simpler process will be furnished by quaternions themselves in the next chapter, and, in giving it, we shall refer back to this section.

85. The associative law of multiplication is now to be proved by means of the distributive (§81). We leave the proof to the student. He has merely to multiply together the factors

$$w + xi + yj + zk, \quad w + x'i + y'j + z'k, \quad \text{and } w'' + x''i + y''j + z''k$$

as follows :

First, multiply the third factor by the second, and then multiply the product by the first; next, multiply the second factor by the first and employ the product to multiply the third: always remembering that the multiplier in any product is placed *before* the multiplicand. He will find the scalar parts and the coefficients of i, j, k, in these products, respectively equal, each to each.

86. With the same expressions for α, β, as in section 83, we have

$$\alpha\beta = (xi + yj + zk)(x^{'}i + y^{'}j + z^{'}k)$$
$$= -(xx^{'} + yy^{'} + zz^{'}) + (yz^{'} - zy^{'})i + (zx^{'} - xz^{'})j + (xy^{'} - yx^{'})k$$

But we have also

$$\beta\alpha = -(xx^{'} + yy^{'} + zz^{'}) - (yz^{'} - zy^{'})i - (zx^{'} - xz^{'})j - (xy^{'} - yx^{'})k$$

The only difference is in the sign of the vector parts. Hence

$$S\alpha\beta = S\beta\alpha \tag{1}$$

$$V\alpha\beta = -V\beta\alpha \tag{2}$$

$$\alpha\beta + \beta\alpha = 2S\alpha\beta \tag{3}$$

$$\alpha\beta - \beta\alpha = 2V\alpha\beta \tag{4}$$

$$\alpha\beta = K.\beta\alpha \tag{5}$$

87. If $\alpha = \beta$ we have of course (§25)

$$x = x^{'}, \quad y = y^{'}, \quad z = z^{'}$$

and the formulae of last section become

$$\alpha\beta = \beta\alpha = \alpha^2 = -(x^2 + y^2 + z^2)$$

which was anticipated in §73, where we proved the formula

$$(T\alpha)^2 = -\alpha^2$$

and also, to a certain extent, in §25.

88. Now let q and r be any quaternions, then

$$\begin{array}{rcl} S.qr & = & S.(Sq + Vq)(Sr + Vr) \\ & = & S.(SqSr + Sr.Vq + Sq.Vr + VqVr) \\ & = & SqSr + S.VqVr \end{array}$$

since the two middle terms are vectors. Similarly,

$$S.rq = SrSq + S.VrVq$$

Hence, since by (1) of §86 we have

$$S.VqVr = S.VrVq$$

we see that

$$S.qr = S.rq \tag{1}$$

a formula of considerable importance.

It may easily be extended to any number of quaternions, because, r being arbitrary, we may put for it rs. Thus we have

$$\begin{aligned} S.qrs &= S.rsq \\ &= S.sqr \end{aligned}$$

by a second application of the process. In words, we have the theorem *the scalar of the product of any number of given quaternions depends only upon the cyclical order in which they are arranged.*

89. An important case is that of three factors, each a vector. The formula then becomes

$$S.\alpha\beta\gamma = S.\beta\gamma\alpha = S.\gamma\alpha\beta$$

But

$$\begin{aligned} S.\alpha\beta\gamma &= S\alpha(S\beta\gamma + V\beta\gamma) \\ &= S\alpha V\beta\gamma && \text{since } \alpha S\beta\gamma \text{ is a vector} \\ &= -S\alpha V\gamma\beta && \text{by (2) of §86} \\ &= -S\alpha(S\gamma\beta + V\gamma\beta) \\ &= -S.\alpha\gamma\beta \end{aligned}$$

Hence *the scalar of the product of three vectors changes sign when the cyclical order is altered.*

By the results of §§55, 73, 79 we see that, for any number of vectors, we have

$$K.\alpha\beta\gamma\ldots\phi\chi = \pm\chi\phi\ldots\gamma\beta\alpha$$

(the positive sign belonging to the product of an even number of vectors) so that

$$S.\alpha\beta\ldots\phi\chi = \pm S.\chi\phi\ldots\beta\alpha$$

Similarly

$$V.\alpha\beta\ldots\phi\chi = \mp V.\chi\phi\ldots\beta\alpha$$

Thus we may generalize (3) and (4) of §86 into

$$2S.\alpha\beta\ldots\phi\chi = \alpha\beta\ldots\chi\phi \pm \phi\chi\ldots\beta\alpha$$

$$2V.\alpha\beta\ldots\phi\chi = \alpha\beta\ldots\chi\phi \mp \phi\chi\ldots\beta\alpha$$

the upper sign still being used when the -number of factors is even.

Other curious propositions connected with this will be given later (some, indeed, will be found in the Examples appended to this chapter), as we wish to develop the really fundamental formulae in as compact a form as possible.

90. By (4) of §86,

$$2V\beta\gamma = \beta\gamma - \gamma\beta$$

Hence

$$2V.\alpha V\beta\gamma = V.\alpha(\beta\gamma - \gamma\beta)$$

(by multiplying both by α, and taking the vector parts of each side)

$$= V(\alpha\beta\gamma + \beta\alpha\gamma - \beta\alpha\gamma - \alpha\gamma\beta)$$

(by introducing the null term $\beta\alpha\gamma - \beta\alpha\gamma$).

That is

$$2V.\alpha V\beta\gamma = V.(\alpha\beta + \beta\alpha)\gamma - V(\beta S\alpha\gamma + \beta V\alpha\gamma + S\alpha\gamma.\beta + V\alpha\gamma.\beta$$

$$= V.(2S\alpha\beta)\gamma - 2V\beta S\alpha\gamma$$

(if we notice that $V(V\alpha\gamma.\beta) = -V.\beta V\alpha\gamma$ by (2) of §86). Hence

$$V.\alpha V\beta\gamma = \gamma S\alpha\beta - \beta S\gamma\alpha \tag{1}$$

a formula of constant occurrence.

Adding $\alpha S\beta\gamma$ to both sides, we get another most valuable formula

$$V.\alpha\beta\gamma = \alpha S\beta\gamma - \beta S\gamma\alpha + \gamma S\alpha\beta \tag{2}$$

and the form of this shows that we may interchange γ and α without altering the right-hand member. This gives

$$V.\alpha\beta\gamma = V.\gamma\beta\alpha$$

a formula which may be greatly extended. (See §89, above.)

Another simple mode of establishing (2) is as follows :

$$\begin{array}{rcl} K.\alpha\beta\gamma & = & -\gamma\beta\alpha \\ \therefore 2V.\alpha\beta\gamma & = & \alpha\beta\gamma - K.\alpha\beta\gamma \text{ (by §79(2))} \\ & = & \alpha\beta\gamma + \gamma\beta\alpha \\ & = & \alpha(\beta\gamma + \gamma\beta) - (\alpha\gamma + \gamma\alpha)\beta + \gamma(\alpha\beta + \beta\alpha) \\ & = & 2\alpha S\beta\gamma - 2\beta S\alpha\gamma + 2\gamma S\alpha\beta \end{array}$$

91. We have also

$$VV\alpha\beta V\gamma\delta = -VV\gamma\delta V\alpha\beta \quad \text{ by (2) of §86}$$

$$= \delta S\gamma V\alpha\beta - \gamma S\delta V\alpha\beta = \delta S.\alpha\beta\gamma - \gamma S.\alpha\beta\delta$$

$$= -\beta S\alpha V\gamma\delta + \alpha S\beta V\gamma\delta = -\beta S.\alpha\gamma\delta + \alpha S.\beta\gamma\delta$$

all of these being arrived at by the help of §90 (1) and of §89; and by treating alternately $V\alpha\beta$ and $V\gamma\delta$ as *simple* vectors.

Equating two of these values, we have

$$\delta S.\alpha\beta\gamma = \alpha S.\beta\gamma\delta + \beta S.\gamma\alpha\delta + \gamma S.\alpha\beta\delta \quad (3$$

a very useful formula, expressing any vector whatever in terms of three given vectors. [This, of course, presupposes that α, β, γ are not coplanar, §23. In fact, if they be coplanar, the factor $S.\alpha\beta\gamma$ vanishes, and thus (3) does not give an expression for δ. This will be shown in §101 below.]

92. That such an expression as (3) is possible we knew already by §23. For variety we may seek another expression of a similar character, by a process which differs entirely from that employed in last section.

α, β, γ being any three non-coplanar vectors, we may derive from them three others $V\alpha\beta$, $V\beta\gamma$, $V\gamma\alpha$ and, as these will not be coplanar, any other vector δ may be expressed as the sum of the three, each multiplied by some scalar. It is required to find this expression for δ.

Let

$$\delta = xV\alpha\beta + yV\beta\gamma + zV\gamma\alpha$$

Then

$$S\gamma\delta = xS.\gamma\alpha\beta = xS.\alpha\beta\gamma$$

the terms in y and z going out, because

$$S\gamma V\beta\gamma = S.\gamma\beta\gamma = S\beta\gamma^2 = \gamma^2 S\beta = 0$$

for γ^2 is (§73) a number.

Similarly

$$S\beta\delta = zS.\beta\gamma\alpha = zS.\alpha\beta\gamma$$

and

$$S\alpha\delta = qS.\alpha\beta\gamma$$

Thus

$$\delta S.\alpha\beta\gamma = V\alpha\beta S\gamma\delta + V\beta\gamma S\alpha\delta + V\gamma\alpha S\beta\delta \quad (4)$$

93. We conclude the chapter by showing (as promised in §64) that the assumption that the product of two parallel vectors is a number, and the product of two perpendicular vectors a third vector perpendicular to both, is not only useful and convenient, but absolutely inevitable, if our system is to deal indifferently with all directions in space. We abridge Hamilton s reasoning.

Suppose that there is no direction in space pre-eminent, and that the product of two vectors is something which has quantity, so as to vary in amount if the factors are changed, and to have its sign changed if that of one of them is reversed ; if the vectors be parallel, their product cannot be, in whole or in part, a vector *inclined* to them, for there is nothing to determine the direction in which it must lie. It cannot be a vector *parallel* to them; for by changing the signs of both factors the product is unchanged, whereas, as the whole system has been reversed, the product vector ought to have been reversed. Hence it must be a number. Again, the product of two perpendicular vectors cannot be wholly or partly a number, because on inverting one of them the sign of that number ought to change; but inverting one of them is simply equivalent to a rotation through two right angles about the other, and (from the symmetry of space) ought to leave the number unchanged. Hence the product of two perpendicular vectors must be a vector, and a simple extension of the same reasoning shows that it must be perpendicular to each of the factors. It is easy to carry this farther, but enough has been said to show the character of the reasoning.

3.5 Examples To Chapter 2.

1. It is obvious from the properties of polar triangles that any mode of representing versors by the *sides* of a spherical triangle must have an equivalent statement in which they are represented by *angles* in the polar triangle.

Show directly that the product of two versors represented by two angles of a spherical triangle is a third versor represented by the *supplement* of the remaining angle of the triangle ; and determine the rule which connects the *directions* in which these angles are to be measured.

2. Hence derive another proof that we have not generally

$$pq = qp$$

3. Hence show that the proof of the associative principle, §57, may be made to depend upon the fact that if from any point of the sphere tangent arcs be drawn to a spherical conic, and also arcs to the foci, the inclination of either tangent arc to one of the focal arcs is equal to that of the other tangent arc to the other focal arc.

4. Prove the formulae

$$2S.\alpha\beta\gamma = \alpha\beta\gamma - \gamma\beta\alpha$$

$$2V.\alpha\beta\gamma = \alpha\beta\gamma + \gamma\beta\alpha$$

5. Show that, whatever odd number of vectors be represented by α, β, γ &c., we have always

$$V.\alpha\beta\gamma\delta\epsilon = V.\epsilon\delta\gamma\beta\alpha$$

$$V.\alpha\beta\gamma\delta\epsilon\zeta\eta = V.\eta\zeta\epsilon\delta\gamma\beta\alpha, \text{ \&c.}$$

6. Show that

$$S.V\alpha\beta V\beta\gamma V\gamma\alpha = -(S.\alpha\beta\gamma)^2$$

$$V.V\alpha\beta V\beta\gamma V\gamma\alpha = V\alpha\beta(\gamma^2 S\alpha\beta - S\beta\gamma S\gamma\alpha) + \dots$$

and

$$V(V\alpha\beta V.V\beta\gamma V\gamma\alpha) = (\beta S\alpha\gamma - \alpha S\beta\gamma)S.\alpha\beta\gamma$$

7. If α, β, γ be any vectors at right angles to each other, show that

$$(\alpha^3 + \beta^3 + \gamma^3)S.\alpha\beta\gamma = \alpha^4 V\beta\gamma + \beta^4 V\gamma\alpha + \gamma^4 V\alpha\beta$$

$$(\alpha^{2n-1} + \beta^{2n-1} + \gamma^{2n-1})S.\alpha\beta\gamma = \alpha^{2n} V\beta\gamma + \beta^{2n} V\gamma\alpha + \gamma^{2n} V\alpha\beta$$

8. If α, β, γ be non-coplanar vectors, find the relations among the six scalars, x, y, z and ξ, η, ζ which are implied in the equation

$$x\alpha + y\beta + z\gamma = \xi V\beta\gamma + \eta V\gamma\alpha + \zeta V\alpha\beta$$

9. If α, β, γ be any three non-coplanar vectors, express any fourth vector, δ, as a linear function of each of the following sets of three derived vectors.

$$V.\gamma\alpha\beta, \quad V.\alpha\beta\gamma, \quad V.\beta\gamma\alpha$$

and

$$V.V\alpha\beta V\beta\gamma V\gamma\alpha, \quad V.V\beta\gamma V\gamma\alpha V\alpha\beta, \quad V.V\gamma\alpha V\alpha\beta V\beta\gamma$$

10. Eliminate ρ from the equations

$$S\alpha\rho = a, \quad S\beta\rho = b, \quad S\gamma\rho = c, \quad S\delta\rho = d$$

where α, β, γ, δ are vectors, and a, b, c, d scalars.

11. In any quadrilateral, plane or gauche, the sum of the squares of the diagonals is double the sum of the squares of the lines joining the middle points of opposite sides.

3.6 Interpretations And Transformations

94. Among the most useful characteristics of the Calculus of Quaternions, the ease of interpreting its formulae geometrically, and the extraordinary variety of transformations of which the simplest expressions are susceptible, deserve a prominent place. We devote this Chapter to some of the more simple of these, together with a few of somewhat more complex character but of constant occurrence in geometrical and physical investigations. Others will appear in every succeeding Chapter. It is here, perhaps, that the student is likely to feel most strongly the peculiar difficulties of the new Calculus. But on that very account he should endeavour to master them, for the variety of forms which any one formula may assume, though puzzling to the beginner, is of the utmost advantage to the advanced student, not alone as aiding him in the solution of complex questions, but as affording an invaluable mental discipline.

95. If we refer again to the figure of §77 we see that

$$OC = OB \cos AOB$$

$$CB = OB \sin AOB$$

Hence if

$$\overline{AB} = \alpha, \quad \overline{OB} = \beta, \quad \text{and } \angle AOB = \theta$$

we have

$$OB = T\beta, \quad OA = T\alpha$$

$$OC = T\beta \cos\theta, \quad CB = T\beta \sin\theta$$

Hence

$$S\frac{\beta}{\alpha} = \frac{OC}{OA} = \frac{T\beta}{T\alpha}\cos\theta$$

Similarly,

$$TV\frac{\beta}{\alpha} = \frac{CB}{OA} = \frac{T\beta}{T\alpha}\sin\theta$$

Hence, if η be a unit-vector perpendicular to α and β, and such that positive rotation about it, through the angle θ, turns α towards β or

$$\eta = \frac{U\overline{CB}}{U\overline{OA}} = U\frac{\overline{CB}}{\overline{OA}} = UV\frac{\beta}{\alpha}$$

we have

$$V\frac{\beta}{\alpha} = \frac{T\beta}{T\alpha}\sin\theta.\eta \qquad \text{(See, again, §84)}$$

96. In the same way, or by putting

$$\begin{aligned} \alpha\beta &= S\alpha\beta + V\alpha\beta \\ &= S\beta\alpha - V\beta\alpha \\ &= \alpha^2\left(S\tfrac{\beta}{\alpha} - V\tfrac{\beta}{\alpha}\right) \\ &= T\alpha^2\left(-S\tfrac{\beta}{\alpha} + V\tfrac{\beta}{\alpha}\right) \end{aligned}$$

we may show that

$$S\alpha\beta = -T\alpha T\beta \cos\theta$$

$$TV\alpha\beta = T\alpha T\beta \sin\theta$$

and

$$V\alpha\beta = T\alpha T\beta \sin\theta.\eta$$

where

$$\eta = UV\alpha\beta = U(-V\beta\alpha) = UV\frac{\beta}{\alpha}$$

Thus *the scalar of the product of two vectors is the continued product of their tensors and of the cosine of the supplement of the contained angle.*

The tensor of the vector of the product of two vectors is the con tinued product of their tensors and the sine of the contained angle ; and the versor of the same is a unit-vector perpendicular to both, and such that the rotation about it from the first vector (i. e. the multiplier) to the second is left-handed or positive.

Hence also $TV\alpha\beta$ is double the area of the triangle two of whose sides are α, β.

97. (a) In any plane triangle ABC we have

$$\overline{AC} = \overline{AB} + \overline{BC}$$

Hence,

$$\overline{AC}^2 = S.\overline{AC}\,\overline{AC} = S.\overline{AC}(\overline{AB} + \overline{BC})$$

With the usual notation for a plane triangle the interpretation of this formula is

$$b^2 = -bc\cos A - ab\cos C$$

or

$$b = c\cos C + c\cos A$$

(b) Again we have, obviously,

$$\begin{aligned} V.\overline{AB}\,\overline{AC} &= V.\overline{AB}(\overline{AB} + \overline{BC}) \\ &= V.\overline{AB}\,\overline{BC} \end{aligned}$$

or

$$cb\sin A = ca\sin B$$

whence

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}$$

These are truths, but not truisms, as we might have been led to fancy from the excessive simplicity of the process employed.

98. From §96 it follows that, if α and β be both actual (i. e. real and non-evanescent) vectors, the equation

$$S\alpha\beta = 0$$

shows that $\cos\theta = 0$, or that α is *perpendicular* to β. And, in fact, we know already that the product of two perpendicular vectors is a vector.

Again if

$$V\alpha\beta = 0$$

we must have $\sin\theta = 0$, or α is *parallel* to β. We know already that the product of two parallel vectors is a scalar.

Hence we see that

$$S\alpha\beta = 0$$

is equivalent to

$$\alpha = V\gamma\beta$$

where γ is an undetermined vector; and that

$$V\alpha\beta = 0$$

is equivalent to

$$\alpha = x\beta$$

where x is an undetermined scalar.

99. If we write, as in §§83, 84,

$$\alpha = ix + jy + kz$$

$$\beta = ix^{'} + jy^{'} + kz^{'}$$

we have, at once, by §86,

$$\begin{array}{rcl} S\alpha\beta & = & -xx^{'} - yy^{'} - zz^{'} \\ & = & -rr^{'}\left(\frac{x}{r}\frac{x^{'}}{r^{'}} + \frac{y}{r}\frac{y^{'}}{r^{'}} + \frac{z}{r}\frac{z^{'}}{r^{'}}\right) \end{array}$$

where

$$r = \sqrt{x^2 + y^2 + z^2}, \quad r^{'} = \sqrt{x^{'2} + y^{'2} + z^{'2}}$$

Also

$$V\alpha\beta = rr^{'}\left\{\frac{yz^{'} - zy^{'}}{rr^{'}}i + \frac{zx^{'} - xz^{'}}{rr^{'}}j + \frac{xy^{'} = yx^{'}}{rr^{'}}k\right\}$$

These express in Cartesian coordinates the propositions we have just proved. In commencing the subject it may perhaps assist the student to see these more familiar forms for the quaternion expressions ; and he will doubtless be induced by their appearance to prosecute the subject, since he cannot fail even at this stage to see how much more simple the quaternion expressions are than those to which he has been accustomed.

100. The expression

$$S.\alpha\beta\gamma$$

may be written

$$SV(\alpha\beta)\gamma$$

because the quaternion $\alpha\beta\gamma$ may be broken up into

$$S(\alpha\beta)\gamma + V(\alpha\beta)\gamma$$

of which the first term is a vector.

But, by §96,

$$SV(\alpha\beta)\gamma = T\alpha T\beta\sin\theta S\eta\gamma$$

Here $T\eta = 1$, let ϕ be the angle between η and γ, then finally

$$S.\alpha\beta\gamma = -T\alpha T\beta T\gamma \sin\theta \cos\phi$$

But as η is perpendicular to α and β, $T\gamma \cos\phi$ is the length of the perpendicular from the extremity of γ upon the plane of α, β. And as the product of the other three factors is (§96) the area of the parallelogram two of whose sides are α, β, we see that the magnitude of $S.\alpha\beta\gamma$, independent of its sign, is *the volume of the parallelepiped of which three coordinate edges are* α, β, γ; or six times the volume of the pyramid which has α, β, γ for edges.

101. Hence the equation

$$S.\alpha\beta\gamma = 0$$

if we suppose $\alpha\beta\gamma$ to be actual vectors, shows either that

$$\sin\theta = 0$$

or

$$\cos\phi = 0$$

i. e. *two of the three vectors are parallel*, or *all three are parallel to one plane*.

This is consistent with previous results, for if $\gamma = p\beta$ we have

$$S.\alpha\beta\gamma = pS.\alpha\beta^2 = 0$$

and, if γ be coplanar with α,β, we have $\gamma = p\alpha + q\beta$ and

$$S.\alpha\beta\gamma = S.\alpha\beta(p\alpha + q\beta) = 0$$

102. This property of the expression $S.\alpha\beta\gamma$ prepares us to find that it is a determinant. And, in fact, if we take α,β as in §83, and in addition

$$\gamma = ix^{''} + jy^{''} + kz^{''}$$

we have at once

$$\begin{aligned} S.\alpha\beta\gamma &= -x^{''}(yz^{'} - zy^{'}) - y^{''}(zx^{'} - xz^{'}) - z^{''}(xy^{'} - yx^{'}) \\ &= -\begin{vmatrix} x & y & z \\ x^{'} & y^{'} & z^{'} \\ x^{''} & y^{''} & z^{''} \end{vmatrix} \end{aligned}$$

The determinant changes sign if we make any two rows change places. This is the proposition we met with before (§89) in the form

$$S.\alpha\beta\gamma = -S.\beta\alpha\gamma = S.\beta\gamma\alpha, \text{ \&c}$$

If we take three new vectors

$$\alpha_1 = ix + jx^{'} + kx^{''}$$

$$\beta_1 = iy + jy' + ky''$$
$$\gamma_1 = iz + jz' + kz''$$

we thus see that they are coplanar if α, β, γ are so. That is, if

$$S.\alpha\beta\gamma = 0$$

then

$$S.\alpha_1\beta_1\gamma_1 = 0$$

103. We have, by §52,

$$\begin{aligned}(Tq)^2 &= qKq = (Sq + Vq)(Sq - Vq) \quad (\S 79)\\ &= (Sq)^2 - (Vq)^2 \quad \text{by algebra}\\ &= (Sq)^2 + (TVq)^2 \quad (\S 73)\end{aligned}$$

If $q = \alpha\beta$, we have $Kq = \beta\alpha$, and the formula becomes

$$\alpha\beta.\beta\alpha = \alpha^2\beta^2 = (S\alpha\beta)^2 - (V\alpha\beta)^2$$

In Cartesian coordinates this is

$$(x^2 + y^2 + z^z)(x'^2 + y'^2 + z'^2)$$
$$= (xx' + yy' + zz')^2 + (yz' - zy')^2 + (zx' - xz')^2 + (xy' - yx')^2$$

More generally we have

$$\begin{aligned}(T(qr))^2 &= (Tq)^2(Tr)^2\\ &= (S.qr)^2 - (V.qr)^2\end{aligned}$$

If we write

$$q = w + \alpha = w + ix + jy + kz$$
$$r = w' + \beta = w' + ix' + jy' + kz'$$

this becomes

$$(w^2 + x^2 + y^2 + z^2)(w'^2 + x'^2 + y'^2 + z'^2)$$
$$= (ww' - xx' - yy' - zz')^2 + (wx' + w'x + yz' - zy')^2$$
$$= (xy' + w'y + zx' - xz')^2 + (wz' + w'z + xy' - yx')^2$$

a formula of algebra due to Euler.

104. We have, of course, by multiplication,

$$(\alpha + \beta)^2 = \alpha^2 + \alpha\beta + \beta\alpha + \beta^2 = \alpha^2 + 2S\alpha\beta + \beta^2 \quad (\S 86\ (3))$$

Translating into the usual notation of plane trigonometry, this becomes

$$c^2 = a^2 - 2ab\cos C + b^2$$

the common formula.

Again,
$$V.(\alpha+\beta)(\alpha-\beta) = -V\alpha\beta + V\beta\alpha = -2V\alpha\beta \qquad (\S 86\ (2)$$

Taking tensors of both sides we have the theorem, *the parallelogram whose sides are parallel and equal to the diagonals of a given parallelogram, has double its area* (§96).

Also
$$S(\alpha+\beta)(\alpha-\beta) = \alpha^2 - \beta^2$$

and vanishes only when $\alpha^2 = \beta^2$, or $T\alpha = T\beta$; that is, *the diagonals of a parallelogram are at right angles to one another, when, and only when, it is a rhombus.*

Later it will be shown that this contains a proof that the angle in a semicircle is a right angle.

105. The expression $\rho = \alpha\beta\alpha^{-1}$
obviously denotes a vector whose tensor is equal to that of β.

But we have $S.\beta\alpha\rho = 0$
so that ρ is in the plane of α, β

Also we have $S\alpha\rho = S\alpha\beta$
so that β and ρ make equal angles with α, evidently on opposite sides of it. Thus if α be the perpendicular to a reflecting surface and β the path of an incident ray, $-\rho$ will be the path of the reflected ray.

Another mode of obtaining these results is to expand the above expression, thus, §90 (2),
$$\begin{aligned}\rho &= 2\alpha^{-1}S\alpha\beta - \beta\\ &= 2\alpha^{-1}S\alpha\beta - \alpha^{-1}(S\alpha\beta + V\alpha\beta)\\ &= \alpha^{-1}(S\alpha\beta - V\alpha\beta)\end{aligned}$$

so that in the figure of §77 we see that if $\overline{OA} = \alpha$, and $\overline{OB} = \beta$, we have $\overline{OD} = \rho = \alpha\beta\alpha^{-1}$

Or, again, we may get the result at once by transforming the equation to $\frac{\rho}{\alpha} = K(\alpha^{-1}\rho) = K\frac{\beta}{\alpha}$

106. For any three coplanar vectors the expression
$$\rho = \alpha\beta\gamma$$

is (§101) a vector. It is interesting to determine what this vector is. The reader will easily see that if a circle be described about the triangle, two of whose sides are (in order) α and β, and if from the extremity of β a line parallel to γ be drawn, again cutting the circle, the vector joining the point of intersection with the origin of α is the direction of the vector $\alpha\beta\gamma$. For we may write it in the form
$$\rho = \alpha\beta^2\beta^{-1}\gamma = -(T\beta)^2\alpha\beta^{-1}\gamma = -(T\beta)^2\frac{\alpha}{\beta}\gamma$$

which shows that the versor $\left(\frac{\alpha}{\beta}\right)$ which turns β into a direction parallel to α, turns γ into a direction parallel to ρ. And this expresses the long-known property of opposite angles of a quadrilateral inscribed in a circle.

Hence if α, β, γ be the sides of a triangle taken in order, the tangents to the circumscribing circle at the angles of the triangle are parallel respectively to

$$\alpha\beta\gamma, \quad \beta\gamma\alpha, \quad \text{and } \gamma\alpha\beta$$

Suppose two of these to be parallel, i. e. let

$$\alpha\beta\gamma = x\beta\gamma\alpha = x\alpha\gamma\beta \quad (\S 90)$$

since the expression is a vector. Hence

$$\beta\gamma = x\gamma\beta$$

which requires either

$$x = 1, \quad V\gamma\beta = 0 \quad \text{or } \gamma||\beta$$

a case not contemplated in the problem; or

$$x = -1, \quad S\beta\gamma = 0$$

i. e. the triangle is right-angled. And geometry shows us at once that this is correct.

Again, if the triangle be isosceles, the tangent at the vertex is parallel to the base. Here we have

$$x\beta = \alpha\beta\gamma$$

or

$$x(\alpha + \gamma) = \alpha(\alpha + \gamma)\gamma$$

whence $x = \gamma^2 = \alpha^2$, or $T\gamma = T\alpha$, as required.

As an elegant extension of this proposition the reader may prove that the vector of the continued product $\alpha\beta\gamma\delta$ of the vectorsides of any quadrilateral inscribed in a sphere is parallel to the radius drawn to the corner (α, δ). [For, if ϵ be the vector from δ, α to β, γ, $\alpha\beta\epsilon$ and $\epsilon\gamma\delta$ are (by what precedes) vectors *touching* the sphere at α, δ. And their product (whose vector part must be parallel to the radius at α, δ) is

$$\alpha\beta\epsilon.\epsilon\gamma\delta = \epsilon^2.\alpha\beta\gamma\delta]$$

107. To exemplify the variety of possible transformations even of simple expressions, we will take cases which are of frequent occurrence in applications to geometry.

Thus

$$T(\rho + \alpha) = T(\rho - \alpha)$$

[which expresses that if

$$\overline{OA} = \alpha \quad \overline{OA'} = -\alpha \quad \text{and} \quad \overline{OP} = \rho$$

we have
$$AP = A'P$$
and thus that P is any point equidistant from two fixed points,] may be written

$$(\rho + \alpha)^2 = (\rho - \alpha)^2$$

or
$$\rho^2 + 2S\alpha\rho + \alpha^2 = \rho^2 - 2S\alpha\rho + \alpha^2 \quad (\S 104)$$
whence
$$S\alpha\rho = 0$$
This may be changed to
$$\alpha\rho + \rho\alpha = 0$$
or
$$\alpha\rho + K\alpha\rho = 0$$
$$SU\frac{\rho}{\alpha} = 0$$
or finally,
$$TVU\frac{\rho}{\alpha} = 1$$
all of which express properties of a plane.

Again,
$$T\rho = T\alpha$$
may be written
$$T\frac{\rho}{\alpha} = 1$$
$$\left(S\frac{\rho}{\alpha}\right)^2 - \left(V\frac{\rho}{\alpha}\right)^2 = 1$$
$$(\rho + \alpha)^2 - 2S\alpha(\rho + \alpha) = 0$$
$$\rho = (\rho + \alpha)^{-1}\alpha(\rho + \alpha)$$
$$S(\rho + \alpha)(\rho - \alpha) = 0$$
or finally,
$$T.(\rho + \alpha)(\rho - \alpha) = 2TV\alpha\rho$$

All of these express properties of a sphere. They will be interpreted when we come to geometrical applications.

108. *To find the space relation among five points.*

A system of five points, so far as its internal relations are concerned, is fully given by the vectors from one to the other four. If three of these be called α, β, γ, the fourth, δ, is necessarily expressible as $x\alpha + y\beta + z\gamma$. Hence the relation required must be independent of x, y, z.

But
$$\left.\begin{array}{rcllll} S\alpha\delta & = & x\alpha^2 & + yS\alpha\beta & + zS\alpha\gamma \\ S\beta\delta & = & xS\beta\alpha & + y\beta^2 & + zS\beta\gamma \\ S\gamma\delta & = & xS\gamma\alpha & + yS\gamma\beta & + z\gamma^2 \\ S\delta\delta = \delta^2 & = & xS\delta\alpha & + yS\delta\beta & + zS\delta\gamma \end{array}\right\} \quad (1)$$

The elimination of x, y, z gives a determinant of the fourth order, which may be written

$$\begin{vmatrix} S\alpha\alpha & S\alpha\beta & S\alpha\gamma & S\alpha\delta \\ S\beta\alpha & S\beta\beta & S\beta\gamma & S\beta\delta \\ S\gamma\alpha & S\gamma\beta & S\gamma\gamma & S\gamma\delta \\ S\delta\alpha & S\delta\beta & S\delta\gamma & S\delta\delta \end{vmatrix} = 0$$

Now each term may be put in either of two forms, thus

$$S\beta\gamma = \frac{1}{2}\left\{\beta^2 + \gamma^2 - (\beta-\gamma)^2\right\} = -T\beta T\gamma\cos\widehat{\beta\gamma}$$

If the former be taken we have the expression connecting the distances, two and two, of five points in the form given by Muir (Proc. R. S. E. 1889) ; if we use the latter, the tensors divide out (some in rows, some in columns), and we have the relation among the cosines of the sides and diagonals of a spherical quadrilateral.

We may easily show (as an exercise in quaternion manipulation merely) that this is the *only* condition, by showing that from it we can get the condition when any other of the points is taken as origin. Thus, let the origin be at α, the vectors are α, $\beta-\alpha$, $\gamma-\alpha$, $\delta-\alpha$. But, by changing the signs of the first row, and first column, of the determinant above, and then adding their values term by term to the other rows and columns, it becomes

$$\begin{vmatrix} S(\quad -\alpha)(-\alpha) & S(\quad -\alpha)(\beta-\alpha) & S(\quad -\alpha)(\gamma-\alpha) & S(\quad -\alpha)(\delta-\alpha) \\ S(\beta-\alpha)(-\alpha) & S(\beta-\alpha)(\beta-\alpha) & S(\beta-\alpha)(\gamma-\alpha) & S(\beta-\alpha)(\delta-\alpha) \\ S(\gamma-\alpha)(-\alpha) & S(\gamma-\alpha)(\beta-\alpha) & S(\gamma-\alpha)(\gamma-\alpha) & S(\gamma-\alpha)(\delta-\alpha) \\ S(\delta-\alpha)(-\alpha) & S(\delta-\alpha)(\beta-\alpha) & S(\delta-\alpha)(\gamma-\alpha) & S(\delta-\alpha)(\delta-\alpha) \end{vmatrix}$$

which, when equated to zero, gives the same relation as before. [See Ex. 10 at the end of this Chapter.]

An additional point, with $\epsilon = x'\alpha + y'\beta + z'\gamma$ gives six additional equations like (1) ; i. e.

$$\begin{array}{rllll} S\alpha\epsilon & = x'\alpha^2 & +y'S\alpha\beta & +z'S\alpha\gamma \\ S\beta\epsilon & = x'S\beta\alpha & +y'\beta^2 & +z'S\beta\gamma \\ S\gamma\epsilon & = x'S\gamma\alpha & +y'S\gamma\beta & +z'\gamma^2 \\ S\delta\epsilon & = x'S\delta\alpha & +y'S\delta\beta & +z'S\delta\gamma \\ & = xS\epsilon\alpha & +yS\epsilon\beta & +zS\epsilon\gamma \\ \epsilon^2 & = x'S\alpha\epsilon & +y'S\beta\epsilon & +z'S\gamma\epsilon \end{array}$$

from which corresponding conclusions may be drawn.

Another mode of solving the problem at the head of this section is to write the *identity*

$$\sum m(\alpha-\theta)^2 = \sum m\alpha^2 - sS.\theta\sum m\alpha + \theta^2\sum m$$

where the ms are undetermined scalars, and the αs are given vectors, while θ is any vector whatever.

Now, *provided that the number of given vectors exceeds four*, we do not completely determine the ms by imposing the conditions

$$\sum m = 0, \quad \sum m\alpha = 0$$

Thus we may write the above identity, for each of five vectors successively, as

$$\begin{array}{rcl}\sum m(\alpha-\alpha_1)^2 & = & \sum m\alpha^2\\ \sum m(\alpha-\alpha_2)^2 & = & \sum m\alpha^2\\ \ldots\ldots & = & \ldots\\ \sum m(\alpha-\alpha_n)^2 & = & \sum m\alpha^2\end{array}$$

Take, with these, $\sum m = 0$
and we have six linear equations from which to eliminate the ms. The resulting determinant is

$$\left|\begin{array}{cccccc}\overline{\alpha_1-\alpha_1^2} & \overline{\alpha_1-\alpha_s^2} & \overline{\alpha_1-\alpha_3^2} & . & \overline{\alpha_1-\alpha_5^2} & 1\\ \overline{\alpha_2-\alpha_1^2} & \overline{\alpha_2-\alpha_s^2} & \overline{\alpha_2-\alpha_3^2} & . & \overline{\alpha_2-\alpha_5^2} & 1\\ . & . & . & & . & \\ . & . & . & & . & \\ \overline{\alpha_5-\alpha_1^2} & \overline{\alpha_5-\alpha_s^2} & \overline{\alpha_5-\alpha_3^2} & . & \overline{\alpha_5-\alpha_5^2} & 1\\ 1 & 1 & . & . & 1 & 0\end{array}\right| \sum m\alpha^2 = 0$$

This is equivalent to the form in which Cayley gave the relation among the mutual distances of five points. (Camb. Math. Journ. 1841.)

109. We have seen in §95 that a quaternion may be divided into its scalar and vector parts as follows:

$$\frac{\beta}{\alpha} = S\frac{\beta}{\alpha} + V\frac{\beta}{\alpha} = \frac{T\beta}{T\alpha}(\cos\theta + \epsilon\sin\theta)$$

where θ is the angle between the directions of α and β and $\epsilon = UV\dfrac{\beta}{\alpha}$ is the unit-vector perpendicular to the plane of α and β so situated that positive (i.e. left-handed) rotation about it turns α towards β

Similarly we have (§96)

$$\begin{array}{rl}\alpha\beta & = S\alpha\beta + V\alpha\beta\\ & = T\alpha T\beta(-\cos\theta + \epsilon\sin\theta)\end{array}$$

θ and ϵ having the same signification as before.

110. Hence, considering the versor parts alone, we have

$$U\frac{\beta}{\alpha} = \cos\theta + \epsilon\sin\theta$$

Similarly

$$U\frac{\gamma}{\beta} = \cos\phi + \epsilon\sin\phi$$

ϕ being the positive angle between the directions of γ and β, and ϵ the same vector as before, if α, β, γ be coplanar.

Also we have

$$U\frac{\gamma}{\alpha} = \cos(\theta+\phi) + \epsilon\sin(\theta+\phi)$$

But we have always

$$\frac{\gamma}{\beta}.\frac{\beta}{\alpha} = \frac{\gamma}{\alpha}$$

and therefore

$$U\frac{\gamma}{\beta}.U\frac{\beta}{\alpha} = U\frac{\gamma}{\alpha}$$

or

$$\begin{aligned}\cos(\phi+\theta) + \epsilon\sin(\phi+\theta) &= (\cos\phi + \epsilon\sin\phi)(\cos\theta + \epsilon\sin\theta)\\ &= \cos\phi\cos\theta - \sin\phi\sin\theta + \epsilon(\sin\phi\cos\theta + \cos\phi\sin\theta)\end{aligned}$$

from which we have at once the fundamental formulae for the cosine and sine of the sum of two arcs, by equating separately the scalar and vector parts of these quaternions.

And we see, as an immediate consequence of the expressions above, that

$$\cos m\theta + \epsilon\sin m\theta = (\cos\theta + \epsilon\sin\theta)^m$$

if m be a positive whole number. For the left-hand side is a versor which turns through the angle $m\theta$ at once, while the right-hand side is a versor which effects the same object by m successive turn ings each through an angle θ. See §§8, 9.

111. To extend this proposition to fractional indices we have only to write $\frac{\theta}{n}$ for θ, when we obtain the results as in ordinary trigonometry.

From De Moivre's Theorem, thus proved, we may of course deduce the rest of Analytical Trigonometry. And as we have already deduced, as interpretations of self-evident quaternion transformations (§§97, 104), the fundamental formulae for the solution of plane triangles, we will now pass to the consideration of spherical trigonometry, a subject specially adapted for treatment by qua ternions; but to which we cannot afford more than a very few sections. (More on this subject will be found in Chap. XI in connexion with the Kinematics of rotation.) The reader is referred to Hamilton s works for the treatment of this subject by quaternion exponentials.

112. Let α, β, γ be unit-vectors drawn from the centr to the corners A, B, C of a triangle on the unit-sphere. Then it is evident that, with the usual notation, we have (§96),

$$S\alpha\beta = -\cos c, \quad S\beta\gamma = -\cos a, \quad S\gamma\alpha = -\cos b$$

$$TV\alpha\beta = \sin c, \quad TV\beta\gamma = \sin a, \quad TV\gamma\alpha = \sin b$$

Also $UV\alpha\beta$, $UV\beta\gamma$, $UV\gamma\alpha$ are evidently the vectors of the corners of the polar triangle.

Hence

$$S.UV\alpha\beta UV\beta\gamma = \cos B, \text{ \&c.}$$

$$TV.UV\alpha\beta UV\beta\gamma = \sin B, \text{ \&c.}$$

Now (§90 (1)) we have

$$\begin{aligned} SV\alpha\beta V\beta\gamma &= S.\alpha V(\beta V\beta\gamma) \\ &= -S\alpha\beta S\beta\gamma + \beta^2 S\alpha\gamma \end{aligned}$$

Remembering that we have

$$SV\alpha\beta V\beta\gamma = TV\alpha\beta TV\beta\gamma S.UV\alpha\beta UV\beta\gamma$$

we see that the formula just written is equivalent to

$$\sin a \sin c \cos B = -\cos a \cos c + \cos b$$

or

$$\cos b = \cos a \cos c + \sin a \sin c \cos B$$

113. Again,

$$V.V\alpha\beta V\beta\gamma = -\beta S\alpha\beta\gamma$$

which gives

$$TV.V\alpha\beta V\beta\gamma = TS.\alpha\beta\gamma = TS.\alpha V\beta\gamma = TS.\beta V\gamma\alpha = TS.\gamma V\alpha\beta$$

or

$$\sin a \sin c \sin B = \sin a \sin p_a = \sin b \sin p_b = \sin c \sin p_c$$

where p_a is the arc drawn from A perpendicular to BC, &c. Hence

$$\sin p_a = \sin c \sin B$$

$$\sin p_b = \frac{\sin a \sin c}{\sin b} \sin B$$

$$\sin p_c = \sin a \sin B$$

114. Combining the results of the last two sections, we have

$$V\alpha\beta.V\beta\gamma = \sin a \sin c \cos B - \beta \sin a \sin c \sin B$$

$$= \sin a \sin c(\cos B - \beta \sin B)$$

$$\left.\begin{array}{ll} \text{Hence} & U.V\alpha\beta V\beta\gamma = (\cos B - \beta \sin B) \\ \text{and} & U.V\gamma\beta V\beta\alpha = (\cos B + \beta \sin B) \end{array}\right\}$$

These are therefore versors which turn all vectors perpendicular to OB negatively or positively about OB through the angle B.

[It will be shown later (§119) that, in the combination

$$(\cos B + \beta \sin B)(\quad)(\cos B - \beta \sin B)$$

the system operated on is made to rotate, as if rigid, round the vector axis β through an angle $2B$.]

As another instance, we have

$$\begin{aligned} \tan B &= \frac{\sin B}{\cos B} \\ &= \frac{TV.V\alpha\beta V\beta\gamma}{S.V\alpha\beta V\beta\gamma} \\ &= -\beta^{-1}\frac{V.V\alpha\beta V\beta\gamma}{S.V\alpha\beta V\beta\gamma} \\ &= -\frac{S.\alpha\beta\gamma}{S\alpha\gamma + S\alpha\beta S\beta\gamma} = \ \&\text{c} \end{aligned} \tag{1}$$

The interpretation of each of these forms gives a different theorem in spherical trigonometry.

115. Again, let us square the equal quantities

$$V.\alpha\beta\gamma \quad \text{and} \quad \alpha S\beta\gamma - \beta S\alpha\gamma + \gamma S\alpha\beta$$

supposing α, β, γ to be any unit-vectors whatever. We have

$$-(V.\alpha\beta\gamma)^2 = S^2\beta\gamma + S^2\gamma\alpha + S^2\alpha\beta + 2S\beta\gamma S\gamma\alpha S\alpha\beta$$

But the left-hand member may be written as

$$T^2.\alpha\beta\gamma - S^2.\alpha\beta\gamma$$

whence

$$1 - S^2.\alpha\beta\gamma = S^2\beta\gamma + S^2\gamma\alpha + S^2\alpha\beta + 2S\beta\gamma S\gamma\alpha S\alpha\beta$$

or

$$\begin{gathered} 1 - \cos^2 a - \cos^2 b - \cos^2 c + 2\cos a \cos b \cos c \\ = \sin^2 a \sin^2 p_a = \ \&\text{c.} \\ = \sin^2 a \sin^2 b \sin^2 C = \ \&\text{c.} \end{gathered}$$

all of which are well-known formulae.

116. Again, for any quaternion,

$$q = Sq + Vq$$

so that, if n be a positive integer,

$$q^n = (Sq)^n + n(Sq)^{n-1}Vq + \frac{n.\overline{n-1}}{1.2}(Sq)^{n-2}(Vq)^2 + \ldots$$

From this at once

$$S.q^n = (Sq)^n - \frac{n.\overline{n-1}}{1.2}(Sq)^{n-2}T^2Vq$$

$$+\frac{n.\overline{n-1}.\overline{n-2}.\overline{n-3}}{1.2.3.4}(Sq)^{n-4}T^4(Vq) - \&c.,$$

$$V.q^n = Vq\left[n(Sq)^{n-1} - \frac{n.\overline{n-1}.\overline{n-2}}{1.2.3}(Sq)^{n-3}T^2Vq + \&c.,\right]$$

If q be a versor we have

$$q = \cos u + \theta \sin u$$

so that

$$\begin{aligned} S.q^n &= (\cos u)^n - \frac{n.\overline{n-1}}{1.2}(\cos u)^{n-2}(\sin u)^2 + \ldots \\ &= \cos nu; \\ V.q^n &= \theta \sin u\left[n(\cos u)^{n-1} - \frac{n.\overline{n-1}.\overline{n-2}}{1.2.3}(\cos u)^{n-3}(\sin u)^2 + \ldots\right] \\ &= \theta \sin nu; \end{aligned}$$

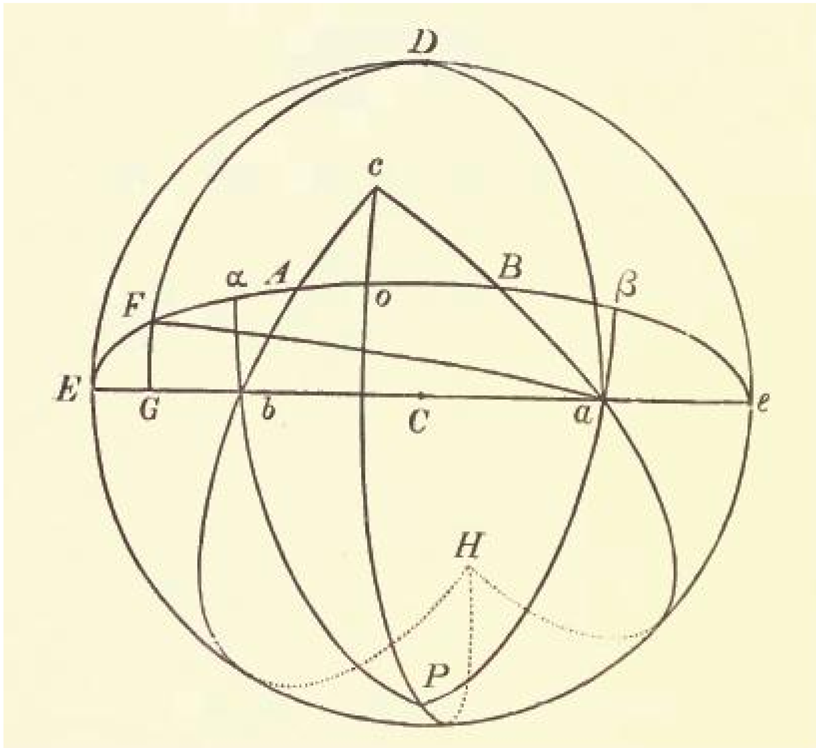
as we might at once have concluded from §110.

Such results may be multiplied indefinitely by any one who has mastered the elements of quaternions.

117. A curious proposition, due to Hamilton, gives us a quaternion expression for the *spherical excess* in any triangle. The following proof, which is very nearly the same as one of his, though by no means the simplest that can be given, is chosen here because it incidentally gives a good deal of other information. We leave the quaternion proof as an exercise.

Let the unit-vectors drawn from the centre of the sphere to A, B, C, respectively, be α, β, γ. It is required to express, as an arc and as an angle on the sphere, the quaternion

$$\beta\alpha^{-1}\gamma$$



The figure represents an orthographic projection made on a plane perpendicular to γ. Hence C is the centre of the circle DEe. Let the great circle through A, B meet DEe in E, e, and let DE be a quadrant. Thus $\overset{\frown}{DE}$ represents γ (§72). Also make $\overset{\frown}{EF} = \overset{\frown}{AB} = \beta\alpha^{-1}$ Then, evidently,

$$\overset{\frown}{DF} = \beta\alpha^{-1}\gamma$$

which gives the arcual representation required.

Let DF cut Ee in G. Make $Ca = EG$, and join D, a, and a, F. Obviously, as D is the pole of Ee, Da is a quadrant ; and since $EG = Ca$, $Ga = EG$, a quadrant also. Hence a is the pole of DG, and therefore the quaternion may be represented by the angle DaF.

Make $Cb = Ca$, and draw the arcs $Pa\beta$, $Pb\alpha$ from P, the pole of AB. Comparing the triangles $Eb\alpha$ and $ea\beta$, we see that $E\alpha = e\beta$. But, since P is the pole of AB, $F\beta a$ is a right angle: and therefore as Fa is a quadrant, so is $F\beta$. Thus AB is the complement of $E\alpha$ or βe, and therefore

$$\alpha\beta = 2AB$$

Join bA. and produce it to c so that $Ac = bA$; join c, P, cutting AB in o. Also join c, B, and B, a.

Since P is the pole of AB, the angles at o are right angles; and therefore, by the equal triangles $b\alpha A$, coA, we have

$$\alpha A = Ao$$

But

$$\alpha\beta = 2AB$$

whence

$$oB = B\beta$$

and therefore the triangles coB and $Ba\beta$ are equal, and c, B, a lie on the same great circle.

Produce cA and cB to meet in H (on the opposite side of the sphere). H and c are diametrically opposite, and therefore cP, produced, passes through H.

Now $Pa = Pb = PH$, for they differ from quadrants by the equal arcs $a\beta$, $b\alpha$, oc. Hence these arcs divide the triangle Hab into three isosceles triangles.

But

$$\angle PHb + \angle PHA = \angle aHb = \angle bca$$

Also

$$\angle Pab = \pi - \angle cab - \angle PaH$$

$$\angle Pba = \angle Pab = \pi - \angle cba - \angle PbH$$

Adding,

$$2\angle Pab = 2\pi - \angle cab - \angle cba - \angle bca$$

$$= \pi - (\text{spherical excess of } abc)$$

But, as $\angle Fa\beta$ and $\angle Dae$ are right angles, we have

$$\text{angle of } \beta\alpha^{-1}\gamma = \angle FaD = \beta ae = \angle Pab$$

$$= \frac{\pi}{2} - \frac{1}{2}(\text{spherical excess of } abc)$$

[Numerous singular geometrical theorems, easily proved *ab initio* by quaternions, follow from this: e.g. The arc AB, which bisects two sides of a spherical triangle abc, intersects the base at the distance of a quadrant from its middle point. All spherical triangles, with a common side, and having their other sides bisected by the same great circle (i.e. having their vertices in a small circle parallel to this great circle) have equal areas, &c.]

118. Let $\overline{Oa} = \alpha^{'}$, $\overline{Ob} = \beta^{'}$, $\overline{Oc} = \gamma^{'}$, and we have

$$\begin{aligned} \left(\frac{\alpha^{'}}{\beta^{'}}\right)^{\frac{1}{2}} \left(\frac{\beta^{'}}{\gamma^{'}}\right)^{\frac{1}{2}} \left(\frac{\gamma^{'}}{\alpha^{'}}\right)^{\frac{1}{2}} &= \frown{Ca}.\frown{cA}.\frown{Bc} \\ &= \frown{Ca}.\frown{BA} \\ &= \frown{EG}.\frown{FE} = \frown{FG} \end{aligned}$$

But FG is the complement of DF. Hence the *angle of the quaternion*

$$\left(\frac{\alpha'}{\beta'}\right)^{\frac{1}{2}}\left(\frac{\beta'}{\gamma'}\right)^{\frac{1}{2}}\left(\frac{\gamma'}{\alpha'}\right)^{\frac{1}{2}}$$

is half the spherical excess of the triangle whose angular points are at the extremities of the unit-vectors α', β', and γ'.

[In seeking a purely quaternion proof of the preceding proposi tions, the student may commence by showing that for any three unit-vectors we have

$$\frac{\beta}{\alpha}\frac{\gamma}{\beta}\frac{\alpha}{\gamma} = -(\beta\alpha^{-1}\gamma)^2$$

The angle of the first of these quaternions can be easily assigned; and the equation shows how to find that of $\beta\alpha^{-1}\gamma$.

Another easy method is to commence afresh by forming from the vectors of the corners of a spherical triangle three new vectors thus:

$$\alpha' = \left(\frac{\beta+\gamma^2}{\alpha}\right)^2 . \alpha, \qquad \&c.$$

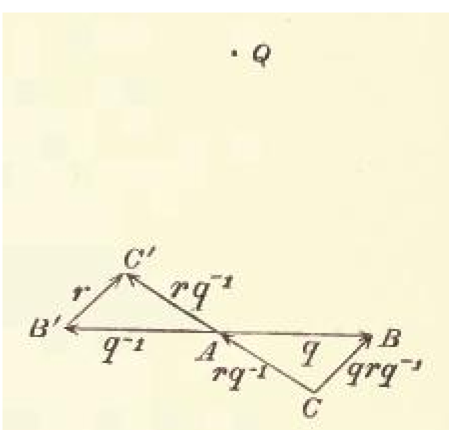
Then the angle between the planes of α, β' and γ', α; or of β, γ' and α', β; or of γ, α' and β', γ is obviously the spherical excess.

But a still simpler method of proof is easily derived from the composition of rotations.]

119. It may be well to introduce here, though it belongs rather to Kinematics than to Geometry, the interpretation of the operator

$$q(\quad)q^{-1}$$

By a rotation, about the axis of q, through double the angle of q, the quaternion r becomes the quaternion qrq^{-1} . Its tensor and angle remain unchanged, its plane or axis alone varies.



A glance at the figure is sufficient for the proof, if we note that of course $T.qrq^{-1} = Tr$, and therefore that we need consider the *versor* parts only. Let Q be the pole of q.

$$\frown{AB} = q, \quad \frown{AB}^{-1} = q^{-1}, \quad \frown{B'C'} = r$$

Join $C'A$, and make $\frown{AC} = \frown{C'A}$. Join CB.

Then $\frown{CB}$ is qrq^{-1}, its arc CB is evidently equal in length to that of r, $B'C'$; and its plane (making the same angle with $B'B$ that that of $B'C'$ does) has evidently been made to revolve about Q, the pole of q, through double the angle of q.

It is obvious, from the nature of the above proof, that this operation is distributive; i.e. that

$$q(r+s)q^{-1} = qrq^{-1} + qsq^{-1}$$

If r be a vector, $= \rho$, then $q\rho q^{-1}$ (which is also a vector) is the result of a rotation through double the angle of q about the axis of q. Hence, as Hamilton has expressed it, if B represent a rigid system, or assemblage of vectors,

$$qBq^{-1}$$

is its new position after rotating through double the angle of q about the axis of q.

120. To compound such rotations, we have

$$r.qBq^{-1}.r^{-1} = rq.B.(rq)^{-1}$$

To cause rotation through an angle t-fold the double of the angle of q we write

$$q^tBq^{-t}$$

To reverse the direction of this rotation write

$$q^{-t}Bq^t$$

To *translate* the body B without rotation, each point of it moving through the vector α, we write $\alpha + B$.

To produce rotation of the translated body about the same axis, and through the same angle, as before,

$$q(\alpha + B)q^{-1}$$

Had we rotated first, and then translated, we should have had

$$\alpha + qBq^{-1}$$

From the point of view of those who do not believe in the Moon s rotation, the former of these expressions ought to be

$$q\alpha q^{-1} + B$$

instead of

$$q\alpha q^{-1} + qBq^{-1}$$

But to such men quaternions are unintelligible.

121. The operator above explained finds, of course, some of its most direct applications in the ordinary questions of Astronomy, connected with the apparent diurnal rotation of the stars. If λ be a unit-vector parallel to the polar axis, and h the hour angle from the meridian, the operator is

$$\left(\cos\frac{h}{2} - \lambda\sin\frac{h}{2}\right)(\quad)\left(\cos\frac{h}{2} + \lambda\sin\frac{h}{2}\right)$$

or

$$L^{-1}(\quad)L$$

the inverse going first, because the *apparent* rotation is negative (clockwise).

If the upward line be i, and the southward j, we have

$$\lambda = i\sin l - j\cos l$$

where l is the latitude of the observer. The meridian equatorial unit vector is

$$\mu = i\cos l + j\sin l$$

and λ, μ, k of course form a rectangular unit system.

The meridian unit-vector of a heavenly body is

$$\delta = i\cos(l-d) + j\sin(l-d)$$

$$= \lambda\sin d + \mu\cos d$$

where d is its declination.

Hence when its hour-angle is h, its vector is

$$\delta' = L^{-1}\delta L$$

The vertical plane containing it intersects the horizon in

$$iVi\delta' = jSj\delta' + kSk\delta'$$

so that

$$\tan(azimuth) = \frac{Sk\delta'}{Sj\delta'} \tag{1}$$

[This may also be obtained directly from the last formula (1) of §114.]

To find its Amplitude, i.e. its azimuth at rising or setting, the hour-angle must be obtained from the condition

$$Si\delta' = 0 \tag{2}$$

These relations, with others immediately deducible from them, enable us (at once and for ever) to dispense with the hideous formulae of Spherical Trigonometry.

122. To show how readily they can be applied, let us translate the expressions above into the ordinary notation. This is effected at once by means of the expressions for λ, μ, L, and δ above, which give by inspection

$$\delta' = \lambda \sin d + (\mu \cos h - k \sin h) \cos d$$

= x sin d + (fjb cos h k sin h) cos d, and we have from (1) and (2) of last section respectively

$$\tan(azimuth) = \frac{\sin h \cos d}{\cos l \sin d - \sin l \cos d \cos h} \tag{1}$$

$$\cos h + \tan l \tan d = 0 \tag{2}$$

In Capt. Weir s ingenious *Azimuth Diagram*, these equations are represented graphically by the rectangular coordinates of a system of confocal conics: viz.

$$\left.\begin{array}{r} x = \sin h \sec l \\ y = \cos h \tan l \end{array}\right\} \tag{3}$$

The ellipses of this system depend upon l alone, the hyperbolas upon h. Since (1) can, by means of (3), be written as

$$\tan(azimuth) = \frac{x}{\tan d - y}$$

we see that the azimuth can be constructed at once by joining with the point $0, -\tan d$, the intersection of the proper ellipse and hyperbola.

Equation (2) puts these expressions for the coordinates in the form

$$\left.\begin{array}{r} x = \sec l\sqrt{1 - \tan^2 l \tan^2 d} \\ y = -\tan^2 l \tan d \end{array}\right\}$$

The elimination of d gives the ellipse as before, but that of l gives, instead of the hyperbolas, the circles

$$x^2 + y^2 - y(\tan d - \cot d) = 1$$

The radius is

$$\frac{1}{2}(\tan d + \cot d)$$

and the coordinates of the centre are

$$0, \quad \frac{1}{2}(\tan d - \cot d)$$

123. A scalar equation in ρ, the vector of an undetermined point, is generally the equation of a *surface*; since we may use in it the expression

$$\rho = x\alpha$$

where x is an unknown scalar, and α any assumed unit-vector. The result is an equation to determine x. Thus one or more points are found on the vector $x\alpha$, whose coordinates satisfy the equation; and the locus is a surface whose degree is determined by that of the equation which gives the values of x.

But a *vector* equation in ρ, as we have seen, generally leads to three scalar equations, from which the three rectangular or other components of the sought vector are to be derived. Such a vector equation, then, usually belongs to a definite number of *points* in space. But in certain cases these may form a *line*, and even a *surface*, the vector equation losing as it were one or two of the three scalar equations to which it is usually equivalent.

Thus while the equation

$$\alpha\rho = \beta$$

gives at once

$$\rho = \alpha^{-1}\beta$$

which is the vector of a definite point, since by making ρ a *vector* we have evidently assumed

$$S\alpha\beta = 0$$

the closely allied equation

$$V\alpha\rho = \beta$$

is easily seen to involve

$$S\alpha\beta = 0$$

and to be satisfied by

$$\rho = \alpha^{-1}\beta + x\alpha$$

whatever be x. Hence the vector of any point whatever in the line drawn parallel to α from the extremity of $\alpha^{-1}\beta$ satisfies the given equation. [The difference between the results depends upon the fact that $S\alpha\rho$ is indeterminate in the second form, but definite ($= 0$) in the first.]

124. Again,

$$V\alpha\rho.V\rho\beta = (V\alpha\beta)^2$$

is equivalent to but two scalar equations. For it shows that $V\alpha\rho$ and $V\beta\rho$ are parallel, i.e. ρ lies in the same plane as α and β, and can therefore be written (§24)

$$\rho = x\alpha + y\beta$$

where x and y are scalars as yet undetermined.

We have now

$$V\alpha\rho = yV\alpha\beta$$

$$V\rho\beta = xV\alpha\beta$$

which, by the given equation, lead to

$$xy = 1, \quad \text{or} \quad y = \frac{1}{x}$$

or finally

$$\rho = x\alpha + \frac{1}{x}\beta$$

which (§40) is the equation of a hyperbola whose asymptotes are in the directions of α and β.

125. Again, the equation

$$V.V\alpha\beta V\alpha\rho = 0$$

though apparently equivalent to three scalar equations, is really equivalent to one only. In fact we see by §91 that it may be written

$$-\alpha S.\alpha\beta\rho = 0$$

whence, if α be not zero, we have

$$S.\alpha\beta\rho = 0$$

and thus (§101) the only condition is that ρ is coplanar with α, β. Hence the equation represents the plane in which α and β lie.

126. Some very curious results are obtained when we extend these processes of interpretation to functions of a *quaternion*

$$q = w + \rho$$

instead of functions of a mere *vector* ρ.

A scalar equation containing such a quaternion, along with quaternion constants, gives, as in last section, the equation of a surface, if we assign a definite value to w. Hence for successive values of w, we have successive surfaces belonging to a system ; and thus when w is indeterminate the equation represents not a *surface*, as before, but a *volume*, in the sense that the vector of any point within that volume satisfies the equation.

Thus the equation

$$(Tq)^2 = a^2$$

or

$$w^2 - \rho^2 = a^2$$

or

$$(TP)^2 = a^2 - w^2$$

represents, for any assigned value of w, not greater than a, a sphere whose radius is $\sqrt{a^2 - w^2}$. Hence the equation is satisfied by the vector of any point whatever in the *volume* of a sphere of radius a, whose centre is origin.

Again, by the same kind of investigation,

$$(T(q - \beta))^2 = a^2$$

where $q = w + \rho$, is easily seen to represent the volume of a sphere of radius a described about the extremity of β as centre.

Also $S(q^2) = -a^2$ is the equation of infinite space less the space contained in a sphere of radius a about the origin.

Similar consequences as to the interpretation of vector equations in quaternions may be readily deduced by the reader.

127. The following transformation is enuntiated without proof by Hamilton (*Lectures*, p. 587, and *Elements*, p. 299).

$$r^{-1}(r^2q^2)^{\frac{1}{2}}q^{-1} = U(rq + KrKq)$$

To prove it, let

$$r^{-1}(r^2q^2)^{\frac{1}{2}}q^{-1} = t$$

then

$$Tt = 1$$

and therefore

$$Kt = t^{-1}$$

But

$$(r^2q^2)^{\frac{1}{2}} = rtq$$

or

$$r^2q^2 = rtqrtq$$

or

$$rq = tqrt$$

Hence

$$KqKr = t^{-1}KrKqt^{-1}$$

or

$$KrKq = tKqKrt$$

Thus we have

$$U(rq \pm KrKq) = tU(qr \pm KqKr)t$$

or, if we put

$$s = U(qr \pm KqKr)$$

$$Ks = \pm tst$$

Hence

$$sKs = (Ts)^2 = 1 = \pm stst$$

which, if we take the positive sign, requires

$$st = \pm 1$$

or

$$t = \pm s^{-1} = \pm UKs$$

which is the required transformation.

[It is to be noticed that there are other results which might have been arrived at by using the negative sign above ; some involving an arbitrary unit-vector, others involving the imaginary of ordinary algebra.]

128. As a final example, we take a transformation of Hamilton's, of great importance in the theory of surfaces of the second order.

Transform the expression

$$(S\alpha\rho)^2 + (S\beta\rho)^2 + (S\gamma\rho)^2$$

in which α, β, γ are any three mutually rectangular vectors, into the form

$$\left(\frac{T(\iota\rho + \rho\kappa)}{\kappa^2 - \iota^2}\right)^2$$

which involves only two vector-constants, ι, κ.

[The student should remark here that ι, κ, two undetermined vectors, involve six disposable constants : and that α, β, γ, being a *rectangular* system, involve also only six constants.]

$$\begin{array}{rcl}\{T(\iota\rho + \rho\kappa)\}^2 & = & (\iota\rho + \rho\kappa)(\rho\iota + \kappa\rho) \quad (\S\S 52, 55) \\ & = & (\iota^2 + \kappa^2)\rho^2 + (\iota\rho\kappa\rho + \rho\kappa\rho\iota) \\ & = & (\iota^2 + \kappa^2)\rho^2 + 2S.\iota\rho\kappa\rho \\ & = & (\iota - \kappa)^2\rho^2 + 4S\iota\rho S\kappa\rho\end{array}$$

Hence

$$(S\alpha\rho)^2 + (S\beta\rho)^2 + (S\gamma\rho)^2 = \frac{(\iota - \kappa)^2}{(\kappa^2 - \iota^2)^2}\rho^2 + 4\frac{S\iota\rho S\kappa\rho}{(\kappa^2 - \iota^2)^2}$$

But

$$\alpha^{-2}(S\alpha\rho)^2 + \beta^{-2}(S\beta\rho)^2 + \gamma^{-2}(S\gamma\rho)^2 = \rho^2 \quad (\S\S 25, 73).$$

Multiply by β^2 and subtract, we get

$$\left(1 - \frac{\beta^2}{\alpha^2}\right)(S\alpha\rho)^2 - \left(\frac{\beta^2}{\gamma^2} - 1\right)(S\gamma\rho)^2 = \left\{\frac{(\iota - \kappa)^2}{(\kappa^2 - \iota^2)^2} - \beta^2\right\}\rho^2 + 4\frac{S\iota\rho S\kappa\rho}{(\kappa^2 - \iota^2)^2}$$

The left side breaks up into two real factors if β^2 be intermediate in value to α^2 and γ^2: and that the right side may do so the term in ρ^2 must vanish. This condition gives

$$\beta^2 = \frac{(\iota-\kappa)^2}{(\kappa^2-\iota^2)^2}$$

and the identity becomes

$$S\left\{\alpha\sqrt{\left(1-\frac{\beta^2}{\alpha^2}\right)}+\gamma\sqrt{\left(\frac{\beta^2}{\gamma^2}-1\right)}\right\}\rho S\left\{\alpha\sqrt{\left(1-\frac{\beta^2}{\alpha^2}\right)}-\gamma\sqrt{\left(\frac{\beta^2}{\gamma^2}-1\right)}\right\}\rho = 4\frac{S\iota\rho S\kappa\rho}{(\kappa^2-\iota^2)^2}$$

Hence we must have

$$\frac{2\iota}{\kappa^2-\iota^2} = p\left\{\alpha\sqrt{\left(1-\frac{\beta^2}{\alpha^2}\right)}+\gamma\sqrt{\left(\frac{\beta^2}{\gamma^2}-1\right)}\right\}$$

$$\frac{2\kappa}{\kappa^2-\iota^2} = \frac{1}{p}\left\{\alpha\sqrt{\left(1-\frac{\beta^2}{\alpha^2}\right)}-\gamma\sqrt{\left(\frac{\beta^2}{\gamma^2}-1\right)}\right\}$$

where ρ is an undetermined scalar.

To determine ρ, substitute in the expression for β^2, and we find

$$\begin{aligned} 4\beta^2 = \tfrac{4(\iota-\kappa)^2}{(\kappa^2-\iota^2)^2} &= \left(p-\tfrac{1}{p}\right)^2(\alpha^2-\beta^2)+\left(p+\tfrac{1}{p}\right)^2(\beta^2-\gamma^2) \\ &= \left(p^2+\tfrac{1}{p^2}\right)(\alpha^2-\gamma^2)-2(\alpha^2+\gamma^2)+4\beta^2 \end{aligned}$$

Thus the transformation succeeds if

$$p^2+\frac{1}{p^2} = \frac{2(\alpha^2+\gamma^2)}{\alpha^2-\gamma^2}$$

which gives

$$p+\frac{1}{p} = \pm 2\sqrt{\frac{\alpha^2}{\alpha^2-\gamma^2}}$$

$$p-\frac{1}{p} = \pm 2\sqrt{\frac{\gamma^2}{\alpha^2-\gamma^2}}$$

Hence

$$\frac{4(\kappa^2-\iota^2)}{(\kappa^2-\iota^2)^2} = \left(\frac{1}{p^2}-p^2\right)(\alpha^2-\gamma^2) = \pm 4\sqrt{\alpha^2\gamma^2}$$

$$(\kappa^2-\iota^2)^{-1} = \pm T\alpha T\gamma$$

Again

$$p = \frac{T\alpha+T\gamma}{\sqrt{\gamma^2-\alpha^2}}, \quad \frac{1}{p} = \frac{T\alpha-T\gamma}{\sqrt{\gamma^2-\alpha^2}}$$

and therefore

$$2\iota = \frac{T\alpha + T\gamma}{T\alpha T\gamma}\left(\sqrt{\frac{\beta^2-\alpha^2}{\gamma^2-\alpha^2}}U\alpha + \sqrt{\frac{\gamma^2-\beta^2}{\gamma^2-\alpha^2}}U\gamma\right)$$

$$2\kappa = \frac{T\alpha - T\gamma}{T\alpha T\gamma}\left(\sqrt{\frac{\beta^2-\alpha^2}{\gamma^2-\alpha^2}}U\alpha - \sqrt{\frac{\gamma^2-\beta^2}{\gamma^2-\alpha^2}}U\gamma\right)$$

Thus we have proved the possibility of the transformation, and determined the transforming vectors ι, κ.

129. By differentiating the equation

$$(S\alpha\rho)^2 + (S\beta\rho)^2 + (S\gamma\rho)^2 = \left(\frac{T(\iota\rho + \rho\kappa)}{(\kappa^2 - \iota^2)}\right)^2$$

we obtain, as will be seen in Chapter IV, the following,

$$S\alpha\rho S\alpha\rho^{'} + S\beta\rho S\beta\rho^{'} + S\gamma\rho S\gamma\rho^{'} = \frac{S.(\iota\rho + \rho\kappa)(\kappa\rho^{'} + \rho^{'}\iota)}{(\kappa^2 - \iota^2)^2}$$

where ρ also may be any vector whatever.

This is another very important formula of transformation ; and it will be a good exercise for the student to prove its truth by processes analogous to those in last section. We may merely observe, what indeed is obvious, that by putting $\rho^{'} = \rho$ it becomes the formula of last section. And we see that we may write, with the recent values of ι and κ in terms of α, β, γ, the identity

$$\begin{aligned} \alpha S\alpha\rho + \beta S\beta\rho + \gamma S\gamma\rho &= \frac{(\iota^2 + \kappa^2)\rho + 2V.\iota\rho\kappa}{(\kappa^2 - \iota^2)^2} \\ &= \frac{(\iota - \kappa)^2\rho + 2(\iota S\kappa\rho + \kappa S\iota\rho)}{(\kappa^2 - \iota^2)^2} \end{aligned}$$

130. In various quaternion investigations, especially in such as involve *imaginary* intersections of curves and surfaces, the old imaginary of algebra of course appears. But it is to be particularly noticed that this expression is analogous to a scalar and not to a vector, and that like real scalars it is commutative in multiplication with all other factors. Thus it appears, by the same proof as in algebra, that any quaternion expression which contains this imaginary can always be broken up into the sum of two parts, one real, the other multiplied by the first power of $\sqrt{-1}$. Such an expression, viz.

$$q = q^{'} + \sqrt{-1}q^{''}$$

where $q^{'}$ and $q^{''}$ are real quaternions, is called by Hamilton a BIQUATERNION. [The student should be warned that the term Biquaternion has since been employed by other writers in the sense sometimes of a "set" of 8 elements, analogous to the Quaternion 4 ; sometimes for an expression $q^{'} + \theta q^{''}$ where θ is not the algebraic imaginary. By them Hamilton s Biquaternion is called simply a quaternion with non-real constituents.] Some little care is requisite in the management of these expressions, but there is no new difficulty. The points to be observed are: first, that any biquaternion can be divided into a real and an imaginary part, the latter being the product of $\sqrt{-1}$ by a real quaternion; second, that this $\sqrt{-1}$ is commutative with all other quantities in multiplication; third, that if two biquaternions be equal, as

$$q^{'} + \sqrt{-1}\, q^{''} = r^{'} + \sqrt{-1}\, r^{''}$$

we have, as in algebra,

$$q^{'} = r^{'}, \quad q^{''} = r^{''}$$

so that an equation between biquaternions involves in general *eight* equations between scalars. Compare §80.

131. We have obviously, since $\sqrt{-1}$ is a scalar,

$$S(q^{'} + \sqrt{-1}\, q^{''}) = Sq^{'} + \sqrt{-1}\, Sq^{''}$$

$$V(q^{'} + \sqrt{-1}\, q^{''}) = Vq^{'} + \sqrt{-1}\, Vq^{''}$$

Hence (§103)

$$\{T(q^{'} + \sqrt{-1}\, q^{''})\}^2$$

$$= (Sq^{'} + \sqrt{-1}\, Sq^{''} + Vq^{'} + \sqrt{-1}\, Vq^{''})(Sq^{'} + \sqrt{-1}\, Sq^{''} - Vq^{'} - \sqrt{-1}\, Vq^{''})$$

$$= (Sq^{'} + \sqrt{-1}\, Sq^{''})^2 - (Vq^{'} + \sqrt{-1}\, Vq^{''})^2$$

$$= (Tq^{'})^2 - (Tq^{''})^2 + 2\sqrt{-1}\, S.q^{'}Kq^{''}$$

The only remark which need be made on such formulae is this, that *the tensor of a biquaternion may vanish while both of the component quaternions are finite.*

Thus, if

$$Tq^{'} = Tq^{''}$$

and

$$S.q^{'}Kq^{''} = 0$$

the above formula gives

$$T(q^{'} + \sqrt{-1}\, q^{''}) = 0$$

The condition

$$S.q^{'}Kq^{''} = 0$$

may be written

$$Kq^{''} = q^{'-1}\alpha, \quad \text{or} \quad q^{''} = -\alpha Kq^{'-1} = -\frac{\alpha q^{'}}{(Tq^{'})^2}$$

where α is any vector whatever.

Hence

$$Tq^{'} = Tq^{''} = TKq^{''} = \frac{T\alpha}{Tq^{''}}$$

and therefore

$$Tq^{'}(Uq^{'} - \sqrt{-1}\, U\alpha.Uq^{'}) = (1 - \sqrt{-1}\, U\alpha)q^{'}$$

is the general form of a biquaternion whose tensor is zero.

132. More generally we have, q, r, $q^{'}$, $r^{'}$ being any four real and non-evanescent quaternions,

$$(q + \sqrt{-1}\, q^{'})(r + \sqrt{-1}\, r^{'}) = qr - q^{'}r^{'} + \sqrt{-1}\,(qr^{'} + q^{'}r)$$

That this product may vanish we must have

$$qr = q^{'}r^{'}$$

and

$$qr^{'} = -q^{'}r$$

Eliminating $r^{'}$ we have

$$qq^{'-1}qr = -q^{'}r$$

which gives

$$(q^{'-1}q)^2 = -1$$

i.e.

$$q = q^{'}\alpha$$

where α is some unit-vector.

And the two equations now agree in giving

$$-r = \alpha r^{'}$$

so that we have the biquaternion factors in the form

$$q^{'}(\alpha + \sqrt{-1}) \quad \text{and} \quad -(\alpha - \sqrt{-1})r^{'}$$

and their product is

$$-q^{'}(\alpha + \sqrt{-1})(\alpha - \sqrt{-1})r^{'}$$

which, of course, vanishes.

[A somewhat simpler investigation of the same proposition may be obtained by writing the biquaternions as

$$q^{'}(q^{'-1}q + \sqrt{-1}) \quad \text{and} \quad (rr^{'-1} + \sqrt{-1})r^{'}$$

or

$$q^{'}(q^{''} + \sqrt{-1}) \quad \text{and} \quad (r^{''} + \sqrt{-1})r^{'}$$

and showing that

$$q^{''} = -r^{''} = \alpha \quad \text{where } T\alpha = 1]$$

From this it appears that if the product of two *bivectors*

$$\rho + \sigma\sqrt{-1} \quad \text{and} \quad \rho^{'} + \sigma^{'}\sqrt{-1}$$

is zero, we must have

$$\sigma^{-1}\rho = -\rho^{'}\sigma^{'-1} = U\alpha$$

where α may be any vector whatever. But this result is still more easily obtained by means of a direct process.

133. It may be well to observe here (as we intend to avail our selves of them in the succeeding Chapters) that certain abbreviated forms of expression may be used when they are not liable to confuse, or lead to error. Thus we may write

$$T^2 q \quad \text{for} \quad (Tq)^2$$

just as we write

$$\cos^2\theta \quad \text{for} \quad (\cos\theta)^2$$

although the true meanings of these expressions are

$$T(Tq) \quad \text{and} \quad \cos(\cos\theta)$$

The former is justifiable, as $T(Tq) = Tq$, and therefore T^2q is not required to signify the second tensor (or tensor of the tensor) of q. But the trigonometrical usage is defensible only on the score of convenience, and is habitually violated by the employment of $cos^{-1}x$ in its natural and proper sense. Similarly we may write

$$S^2 q \quad \text{for} \quad (Sq)^2, \quad \&c.$$

but it may be advisable not to use

$$Sq^2$$

as the equivalent of either of those just written; inasmuch as it might be confounded with the (generally) different quantity

$$S.q^2 \quad \text{or} \quad S(q^2)$$

although this is rarely written without the point or the brackets.

The question of the use of points or brackets is one on which no very definite rules can be laid down. A beginner ought to use them freely, and he will soon learn by trial which of them are absolutely necessary to prevent ambiguity.

In the present work this course has been adopted:– the earlier examples in each part of the subject being treated with a free use of points and brackets, while in the later examples superfluous marks of the kind are gradually got rid of.

It may be well to indicate some general principles which regulate the omission of these marks. Thus in $S.\alpha\beta$ or $V.\alpha\beta$ the point is obviously unnecessary:– because $S\alpha = 0$, and $V\alpha = \alpha$ so that the S would annihilate the term if it applied to α alone, while in the same case the V would be superfluous. But in $S.qr$ and $V.qr$, the point (or an equivalent) is indispensable, for $Sq.r$, and $Vq.r$ are usually quite different from the first written quantities. In the case of K, and of d (used for scalar differentiation), the *omission* of the point indicates that the operator acts *only* on the nearest factor:– thus

$$Kqr = (Kq)r = Kq.r, \quad dqr = (dq)r = dq.r$$

Kqr = (Kq) r = Kq . r, dqr = (dq) r=dq.r; while, if its action extend farther, we write

$$K.qr = K(qr), \quad d.qr = d(qr) \quad \&c.$$

In more complex cases we must be ruled by the general principle of dropping nothing which is essential. Thus, for instance

$$V(pK(dq)V(Vq.r))$$

may be written without ambiguity as

$$V(pK(dq)V(Vq.r))$$

but nothing more can be dropped without altering its value.

Another peculiarity of notation, which will occasionally be required, shows *which portions* of a complex product are affected by an operator. Thus we write

$$\nabla S\sigma\tau$$

if ∇ operates on σ and also on τ, but

$$\nabla_1 S\sigma\tau_1$$

if it operates on τ alone. See, in this connection, the last Example at the end of Chap. IV. below.

134. The beginner may expect to be at first a little puzzled with this aspect of the notation; but, as he learns more of the subject, he will soon see clearly the distinction between such an expression as

$$S.V\alpha\beta V\beta\gamma$$

where we may omit at pleasure either the point or the first V without altering the value, and the very different one

$$S\alpha\beta.V\beta\gamma$$

which admits of no such changes, without alteration of its value.

All these simplifications of notation are, in fact, merely examples of the transformations of quaternion expressions to which part of this Chapter has been devoted. Thus, to take a very simple ex ample, we easily see that

$$\begin{aligned} S.V\alpha\beta V\beta\gamma &= SV\alpha\beta V\beta\gamma = S.\alpha\beta V\beta\gamma = S\alpha V.\beta V\beta\gamma = -S\alpha V.(V\beta\gamma)\beta \\ &= S\alpha V.(V\gamma\beta)\beta = S.\alpha V(\gamma\beta)\beta = S.V(\gamma\beta)\beta\alpha = SV\gamma\beta V\beta\alpha \\ &= S.\gamma\beta V\beta\alpha = S.K(\beta\gamma)V\beta\alpha = S.\beta\gamma KV\beta\alpha = -S.\beta\gamma V\beta\alpha \\ &= S.V\gamma\beta V\beta\alpha, \&c., \&c. \end{aligned}$$

The above group does not nearly exhaust the list of even the simpler ways of expressing the given quantity. We recommend it to the careful study of the reader. He will find it advisable, at first, to use stops and brackets pretty freely; but will gradually learn to dispense with those which are not absolutely necessary to prevent ambiguity.

There is, however, one additional point of notation to which the reader s attention should be most carefully directed. A very simple instance will suffice. Take the expressions

$$\frac{\beta}{\gamma}.\frac{\gamma}{\alpha} \quad \text{and} \quad \frac{\beta\gamma}{\gamma\alpha}$$

The first of these is

$$\beta\gamma^{-1}.\gamma\alpha^{-1} = \beta\alpha^{-1}$$

and presents no difficulty. But the second, though at first sight it closely resembles the first, is in general totally different in value, being in fact equal to

$$\beta\gamma\alpha^{-1}\gamma^{-1}$$

For the denominator must be treated as *one quaternion*. If, then, we write

$$\frac{\beta\gamma}{\gamma\alpha} = q$$

we have

$$\beta\gamma = q\gamma\alpha$$

so that, as stated above,

$$q = \beta\gamma\alpha^{-1}\gamma^{-1}$$

We see therefore that

$$\frac{\beta}{\gamma}.\frac{\gamma}{\alpha} = \frac{\beta}{\alpha} = \frac{\beta\gamma}{\alpha\gamma}; \quad \text{but } \textit{not} \quad = \frac{\beta\gamma}{\gamma\alpha}$$

3.7 Examples to Chapter 3

1. Investigate, by quaternions, the requisite formulae for changing from any one set of coordinate axes to another ; and derive from your general result, and also from special investiga tions, the usual expressions for the following cases:

(a) Rectangular axes turned about z through any angle.

(b) Rectangular axes turned into any new position by rota tion about a line equally inclined to the three.

(c) Rectangular turned to oblique, one of the new axes lying in each of the former coordinate planes.

2. Point out the distinction between

$$\left(\frac{\alpha+\beta}{\alpha}\right)^2 \quad \text{and} \quad \frac{(\alpha+\beta)^2}{\alpha^2}$$

and find the value of their difference.

If

$$T\beta/\alpha = 1 \quad \text{and} \quad U\frac{\alpha+\beta}{\alpha} = \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2}}$$

Show also that

$$\frac{\alpha+\beta}{\alpha-\beta} = \frac{V\alpha\beta}{1+S\alpha\beta'}$$

and

$$\frac{\alpha-\beta}{\alpha+\beta} = -\frac{V\alpha\beta}{1-S\alpha\beta'}$$

provided α and β be unit-vectors. If these conditions are not fulfilled, what are the true values ?

3. Show that, whatever quaternion r may be, the expression

$$\alpha r + r\beta$$

in which α and β are any two unit- vectors, is reducible to the form

$$l(\alpha+\beta) + m(\alpha\beta - 1)$$

where l and m are scalars.

4. If $Tp = T\alpha = T\beta = 1$, and $S.\alpha\beta\rho = 0$ show by direct transformations that

$$S.U(\rho-\alpha)U(\rho-\beta) = \pm\sqrt{\frac{1}{2}(1 - S\alpha\beta)}$$

Interpret this theorem geometrically.

5. If $S\alpha\beta = 0$, $T\alpha = T\beta = 1$, show that

$$(1+\alpha^m)\beta = 2\cos\frac{m\pi}{4}\alpha^{\frac{m}{2}}\beta = 2S\alpha^{\frac{m}{2}}.\alpha^{\frac{m}{2}}\beta$$

6. Put in its simplest form the equation

$$\rho S.V\alpha\beta V\beta\gamma V\gamma\alpha = aV.V\gamma\alpha V\alpha\beta + bV.V\alpha\beta V\beta\gamma + cV.V\beta\gamma V\gamma\alpha$$

and show that

$$a = S.\beta\gamma\rho, \quad \&c.$$

7. Show that any quaternion may in general, in one way only, be expressed as a homogeneous linear function of four given quaternions. Point out the nature of the exceptional cases. Also find the simplest form in which any quaternion may generally be expressed in terms of two given quaternions.

8. Prove the following theorems, and exhibit them as properties of determinants :

(a) $S.(\alpha+\beta)(\beta+\gamma)(\gamma+\alpha) = 2S.\alpha\beta\gamma$

(b) $S.V\alpha\beta V\beta\gamma V\gamma\alpha = -(S.\alpha\beta\gamma)^2$

(c) $S.V(\alpha+\beta)(\beta+\gamma)V(\beta+\gamma)(\gamma+\alpha)V(\gamma+\alpha)(\alpha+\beta) = -4(S.\alpha\beta\gamma)^2$

(d) $S.V(V\alpha\beta V\beta\gamma)V(V\beta\gamma V\gamma\alpha)V(V\gamma\alpha V\alpha\beta) = -(S.\alpha\beta\gamma)^4$

(e) $S.\delta\epsilon\zeta = -16(S.\alpha\beta\gamma)^4$
where

$$\delta = V(V(\alpha+\beta)(\beta+\gamma)V(\beta+\gamma)(\gamma+\alpha))$$
$$\epsilon = V(V(\beta+\gamma)(\gamma+\alpha)V(\gamma+\alpha)(\alpha+\beta))$$
$$\zeta = V(V(\gamma+\alpha)(\alpha+\beta)V(\alpha+\beta)(\beta+\gamma))$$

9. Prove the common formula for the product of two determinants of the third order in the form

$$S.\alpha\beta\gamma S.\alpha_1\beta_1\gamma_1 = \begin{vmatrix} S\alpha\alpha_1 & S\beta\alpha_1 & S\gamma\alpha_1 \\ S\alpha\beta_1 & S\beta\beta_1 & S\gamma\beta_1 \\ S\alpha\gamma_1 & S\beta\gamma_1 & S\gamma\gamma_1 \end{vmatrix}$$

10. Show that, whatever be the eight vectors involved,

$$\begin{vmatrix} S\alpha\alpha_1 & S\alpha\beta_1 & S\alpha\gamma_1 & S\alpha\delta_1 \\ S\beta\alpha_1 & S\beta\beta_1 & S\beta\gamma_1 & S\beta\delta_1 \\ S\gamma\alpha_1 & S\gamma\beta_1 & S\gamma\gamma_1 & S\gamma\delta_1 \\ S\delta\alpha_1 & S\delta\beta_1 & S\delta\gamma_1 & S\delta\delta_1 \end{vmatrix} = S.\alpha\beta\gamma S.\beta_1\gamma_1\delta_1 S\alpha_1(\delta-\delta) = 0$$

If the single term $S\alpha\alpha_1$, be changed to $S\alpha_0\alpha_1$, the value of the determinant is

$$S.\beta\gamma\delta S.\beta_1\gamma_1\delta_1 S\alpha_1(\alpha_0-\alpha)$$

State these as propositions in spherical trigonometry.

Form the corresponding null determinant for any two groups of five quaternions : and give its geometrical interpretation.

11. If, in §102, α, β, γ be three mutually perpendicular vectors, can anything be predicated as to α_1, β_1, γ_1? If α, β, γ be rectangular unit-vectors, what of α_1, β_1, γ_1?

12. If α, β, γ, $\alpha^{'}$, $\beta^{'}$, $\gamma^{'}$ be two sets of rectangular unit-vectors, show that

$$S\alpha\alpha^{'} = S\gamma\beta^{'}S\beta\gamma^{'} = S\beta\beta^{'}S\gamma\gamma^{'} \quad \text{\&c. \&c.}$$

13. The lines bisecting pairs of opposite sides of a quadrilateral (plane or gauche) are perpendicular to each other when the diagonals of the quadrilateral are equal.

14. Show that

(a) $S.q^2 = 2S^2q - T^2q$

(b) $S.q^3 = S^3q - 3SqT^2Vq$

(c) $\alpha^2\beta^2\gamma^2 + S^2.\alpha\beta\gamma = V^2.\alpha\beta\gamma$

(d) $S(V.\alpha\beta\gamma V.\beta\gamma\alpha V.\gamma\alpha\beta) = 4S\alpha\beta S\beta\gamma S\gamma\alpha S.\alpha\beta\gamma$

(e) $V.q^3 = (2S^2q - T^2Vq)Vq$

(f) $qUVq^{-1} = -Sq.UVq + TVq$

and interpret each as a formula in plane or spherical trigonometry.

15. If q be an undetermined quaternion, what loci are represented by

(a) $(q\alpha^{-1})^2 = -a^2$

(b) $(q\alpha^{-1})^4 = a^4$

(c) $S.(q - \alpha)^2 = a^2$

where a is any given scalar and α any given vector ?

16. If q be any quaternion, show that the equation

$$Q^2 = q^2$$

is satisfied, not alone by $Q = \pm q$, but also by

$$Q = \pm\sqrt{-1}(Sq.UVq - TVq)$$

(Hamilton, *Lectures*, p. 673.)

17. Wherein consists the difference between the two equations

$$T^2\frac{\rho}{\alpha} = 1 \quad \text{and} \quad \left(\frac{\rho}{\alpha}\right)^2 = -1$$

What is the full interpretation of each, α being a given, and p an undetermined, vector?

18. Find the *full* consequences of each of the following groups of equations, as regards both the unknown vector ρ and the given vectors α, β, γ:

$$(a)\ \begin{array}{r} S.\alpha\beta\rho = 0 \\ \\ S.\beta\gamma\rho = 0 \end{array} \qquad (b)\ \begin{array}{r} S\alpha\rho = 0 \\ S.\alpha\beta\rho = 0 \\ S\beta\rho = 0 \end{array} \qquad (c)\ \begin{array}{r} S\alpha\rho = 0 \\ S.\alpha\beta\rho = 0 \\ S.\alpha\beta\gamma\rho = 0 \end{array}$$

19. From §§74, 110, show that, if ϵ be any unit-vector, and m any scalar,

$$\epsilon^m = \cos\frac{m\pi}{2} + \epsilon\sin\frac{m\pi}{2}$$

Hence show that if α, β, γ be radii drawn to the corners of a triangle on the unit-sphere, whose spherical excess is m right angles,

$$\frac{\alpha+\beta}{\beta+\gamma}.\frac{\gamma+\alpha}{\alpha+\beta}.\frac{\beta+\gamma}{\gamma+\alpha} = \alpha^m$$

Also that, if A, B, C be the angles of the triangle, we have

$$\gamma^{\frac{2C}{\pi}}\beta^{\frac{2B}{\pi}}\alpha^{\frac{2A}{\pi}} = -1$$

20. Show that for any three vectors α, β, γ we have

$$(U\alpha\beta)^2 + (U\beta\gamma)^2 + (U\alpha\gamma)^2 + (U.\alpha\beta\gamma)^2 + 4U\alpha\gamma.SU\alpha\beta SU\beta\gamma = -2$$

(Hamilton, *Elements*, p. 388.)

21. If a_1, a_2, a_3, x be any four scalars, and ρ_1, ρ_2, ρ_3 any three vectors, show that

$$(S.\rho_1\rho_2\rho_3)^2 + (\sum.a_1V\rho_2\rho_3)^2 + x^2(\sum V\rho_1\rho_2)^2 -$$

$$x^2(\sum.a_1(\rho_2-\rho_3))^2 + 2\prod(x^2 + S\rho_1\rho_2 + a_1a_2)$$

$$= 2\prod(x^2+\rho^2) + 2\prod a^2 +$$

$$\sum\{(x^2+a_1^2+\rho_1^2)((V\rho_2\rho_3)^2 + 2a_2a_3(x^2+S\rho_2\rho_3) - x^2(\rho_2-\rho_3)^2)\}$$

where $\prod a^2 = a_1^2a_2^2a_3^2$

Verify this formula by a simple process in the particular case

$$a_1 = a_2 = a_3 = x = 0$$

(*Ibid*)

22. Eliminate p from the equations

$$V.\beta\rho\alpha\rho = 0, \quad S\gamma\rho = 0$$

and state the problem and its solution in a geometrical form.

23. If p, q, r, s be four versors, such that

$$qp = -sr = \alpha$$

$$rq = -ps = \beta$$

where α and β are unit-vectors; show that

$$S(V.VsVqV.VrVp) = 0$$

Interpret this as a property of a spherical quadrilateral.

24. Show that, if pq, rs, pr, and qs be vectors, we have

$$S(V.VpVsV.VqVr) = 0$$

25. If α, β, γ be unit-vectors,

$$V\beta\gamma S.\alpha\beta\gamma = -\alpha(1 - S^2\beta\gamma) - \beta(S\alpha\gamma S\beta r + S\alpha\beta) - \gamma(S\alpha\beta S\beta\gamma + S\alpha\gamma)$$

26. If i, j, k, $i^{'}$, $j^{'}$, $k^{'}$, be two sets of rectangular unit-vectors, show that

$$\begin{aligned} S.Vii^{'}Vjj^{'}Vkk^{'} &= (Sij^{'})^2 - (Sji^{'})^2 \\ &= (Sjk^{'})^2 - (Skj^{'})^2 = \&\text{c.} \end{aligned}$$

and find the values of the vector of the same product.

27. If α, β, γ be a rectangular unit-vector system, show that, whatever be λ, μ, ν

$$\lambda S^2 i\alpha + \mu S^2 j\gamma + \nu S^2 k\beta$$

$$\lambda S^2 k\gamma + \mu S^2 i\beta + \nu S^2 j\alpha$$

and

$$\lambda S^2 j\beta + \mu S^2 k\alpha + \nu S^2 i\gamma$$

are coplanar vectors. What is the connection between this and the result of the preceding example ?

3.8 Axiom Examples

The basic operation for creating quaternions is **quatern**. This is a quaternion over the rational numbers.

```
q:=quatern(2/11,-8,3/4,1)
```

$$\frac{2}{11} - 8\ i + \frac{3}{4}\ j + k$$

`Type: Quaternion Fraction Integer`

This is a quaternion over the integers.

```
r:=quatern(1,2,3,4)
```

$$1 + 2\ i + 3\ j + 4\ k$$

`Type: Quaternion Integer`

We can also construct quaternions with complex components. First we construct a complex number.

```
b:=complex(3,4)
```

$$3 + 4\ i$$

`Type: Complex Integer`

and then we use it as a component in a quaternion.

```
s:=quatern(3,1/7,b,2)
```

$$3 + \frac{1}{7}\ i + (3 + 4\ i)\ j + 2\ k$$

`Type: Quaternion Complex Fraction Integer`

Notice that the i component of the complex number has no relation to the i component of the quaternion even though they use the same symbol by convention.

The four parts of a quaternion are the real part, the i imaginary part, the j imaginary part, and the k imaginary part. The **real** function returns the real part.

```
real q
```

$$\frac{2}{11}$$

Type: Fraction Integer

The **imagI** function returns the i imaginary part.

```
imagI q
```

$$-8$$

Type: Fraction Integer

The **imagJ** function returns the j imaginary part.

```
imagJ q
```

$$\frac{3}{4}$$

Type: Fraction Integer

The **imagK** function returns the k imaginary part.

```
imagK q
```

$$1$$

Type: Fraction Integer

Quaternions satisfy a very fundamental relationship between the parts, namely that

$$i^2 = j^2 = k^2 = ijk = -1$$

. This is similar to the requirement in complex numbers of the form $a + bi$ that $i^2 = -1$.

The set of quaternions is denoted by $\mathbb{H}$, whereas the integers are denoted by $\mathbb{Z}$ and the complex numbers by $\mathbb{C}$.

Quaternions are not commutative which means that in general

$$AB \neq BA$$

for any two quaternions, A and B. So, for instance,

```
q*r
```

$$\frac{437}{44} - \frac{84}{11}\ i + \frac{1553}{44}\ j - \frac{523}{22}\ k$$

Type: Quaternion Fraction Integer

```
r*q
```

$$\frac{437}{44} - \frac{84}{11}\ i - \frac{1439}{44}\ j + \frac{599}{22}\ k$$

Type: Quaternion Fraction Integer

and these are clearly not equal.

Complex 2×2 matrices form an alternate, equivalent representation of quaternions. These matrices have the form:

$$\left[\begin{array}{cc} u & v \\ -\overline{v} & \overline{u} \end{array}\right]$$

=

$$\left[\begin{array}{cc} a+bi & c+di \\ -c+di & a-bi \end{array}\right]$$

where u and v are complex, $\overline{u}$ is complex conjugate of u, $\overline{z}$ is the complex conjugate of z, and a,b,c, and d are real.

Within the quaternion each component operator represents a basis element in $\mathbb{R}^4$ thus:

$$1 = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$i = \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{array}\right]$$

Chapter 4

Clifford Algebra[Fl09]

This is quoted from John Fletcher's web page[Fl09] (with permission).

The theory of Clifford Algebra includes a statement that each Clifford Algebra is isomorphic to a matrix representation. Several authors discuss this and in particular Ablamowicz[Ab98] gives examples of derivation of the matrix representation. A matrix will itself satisfy the characteristic polynomial equation obeyed by its own eigenvalues. This relationship can be used to calculate the inverse of a matrix from powers of the matrix itself. It is demonstrated that the matrix basis of a Clifford number can be used to calculate the inverse of a Clifford number using the characteristic equation of the matrix and powers of the Clifford number. Examples are given for the algebras Clifford(2), Clifford(3) and Clifford(2,2).

4.1 Introduction

Introductory texts on Clifford algebra state that for any chosen Clifford Algebra there is a matrix representation which is equivalent. Several authors discuss this in more detail and in particular, Ablamowicz[Ab98] shows that the matrices can be derived for each algebra from a choice of idempotent, a member of the algebra which when squared gives itself. The idea of this paper is that any matrix obeys the characteristic equation of its own eigenvalues, and that therefore the equivalent Clifford number will also obey the same characteristic equation. This relationship can be exploited to calculate the inverse of a Clifford number. This result can be used symbolically to find the general form of the inverse in a particular algebra, and also in numerical work to calculate the inverse of a particular member. This latter approach needs the knowledge of the matrices. Ablamowicz has provided a method for generating them in the form of a Maple implementation. This knowledge is not believed to be new, but the theory is distributed in the literature and the purpose of this paper is to make it clear.

The examples have been first developed using a system of symbolic algebra described in another paper by this author[Fl01].

4.2 Clifford Basis Matrix Theory

The theory of the matrix basis is discussed extensively by Ablamowicz. This theory will be illustrated here following the notation of Ablamowicz by reference to Clifford(2) algebra and can be applied to other Clifford Algebras. For most Clifford algebras there is at least one primitive idempotent, such that it squares to itself. For Clifford (2), which has two basis members e_1 and e_2, one such idempotent involves only one of the basis members, e_1, i.e.

$$f_1 = f = \frac{1}{2}(1 + e_1)$$

If the idempotent is mutiplied by the other basis function e_2, other functions can be generated:

$$f_2 = e_2 f = \left(\frac{1}{2} - \frac{1}{2}e_1\right) e_2$$

$$f_3 = f e_2 = \left(\frac{1}{2} + \frac{1}{2}e_1\right) e_2$$

$$f_4 = e_2 f e_2 = \frac{1}{2} - \frac{1}{2}e_1$$

Note that $fe_2 2f = 0$. These four functions provide a means of representing any member of the space, so that if a general member c is given in terms of the basis members of the algebra

$$c = a_0 + a_1 e_1 + a_2 e_2 + a_3 e_1 e_2$$

it can also be represented by a series of terms in the idempotent and the other functions.

$$\begin{aligned} c &= a_{11} f_1 + a_{21} f_2 + a_{12} f_3 + a_{22} f_4 \\ &= \tfrac{1}{2} a_{11} + \tfrac{1}{2} a_{11} e_1 + \tfrac{1}{2} a_{21} e_2 - \tfrac{1}{2} a_{21} e_1 e_2 + \\ &\quad \tfrac{1}{2} a_{12} e_2 + \tfrac{1}{2} a_{12} e_1 e_2 + \tfrac{1}{2} a_{22} - \tfrac{1}{2} a_{22} e_1 \end{aligned}$$

Equating coefficients it is clear that the following equations apply.

$$\begin{aligned} a_0 &= \tfrac{1}{2}a_{11} + \tfrac{1}{2}a_{22} \\ a_1 &= \tfrac{1}{2}a_{11} - \tfrac{1}{2}a_{22} \\ a_2 &= \tfrac{1}{2}a_{12} + \tfrac{1}{2}a_{21} \\ a_3 &= \tfrac{1}{2}a_{12} - \tfrac{1}{2}a_{21} \end{aligned}$$

The reverse equations can be recovered by multiplying the two forms of c by different combinations of the functions f_1, f_2 and f_3. The equation

$$\begin{aligned} f_1 c f_1 &= f_1(a_{11}f_1 + a_{21}f_2 + a_{12}f_3 + a_{22}f_4)f_1 \\ &= f_1(a_0 + a_1e_1 + a_2e_2 + a_3e_1e_2)f_1 \end{aligned}$$

reduces to the equation

$$a_{11}f = (a_0 + a_1)f$$

and similar equations can be deduced from other combinations of the functions as follows.

$$\begin{aligned} f_1cf_2 : a_{12}f &= (a_2 + a_3)f \\ f_2cf_1 : a_{21}f &= (a_2 - a_3)f \\ f_3cf_2 : a_{22}f &= (a_0 - a_1)f \end{aligned}$$

If a matrix is defined as

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

so that

$$Af = \begin{pmatrix} a_{11}f & a_{12}f \\ a_{21}f & a_{22}f \end{pmatrix} = \begin{pmatrix} a_0 + a_1 & a_2 + a_3 \\ a_2 - a_3 & a_0 - a_1 \end{pmatrix} f$$

then the expression

$$\begin{pmatrix} 1 & e_2 \end{pmatrix} \begin{pmatrix} a_{11}f & a_{12}f \\ a_{21}f & a_{22}f \end{pmatrix} \begin{pmatrix} 1 \\ e_2 \end{pmatrix} = a_{11}f_1 + a_{21}f_2 + a_{12}f_3 + a_{22}f_4 = c$$

generates the general Clifford object c. All that remains to form the basis matrices is to make c each basis member in turn, and named as shown.

$$
\begin{array}{llcll}
c = 1: & Af & = & \begin{pmatrix} f & 0 \\ 0 & f \end{pmatrix} & = & E_0 f \\
c = e_1 & Af & = & \begin{pmatrix} f & 0 \\ 0 & -f \end{pmatrix} & = & E_1 f \\
c = e_2 & Af & = & \begin{pmatrix} 0 & f \\ f & 0 \end{pmatrix} & = & E_2 f \\
c = e_1 e_2 & Af & = & \begin{pmatrix} 0 & f \\ -f & 0 \end{pmatrix} & = & E_{12} f
\end{array}
$$

These are the usual basis matrices for Clifford (2) except that they are multiplied by the idempotent.

This approach provides an explanation for the basis matrices in terms only of the Clifford Algebra itself. They are the matrix representation of the basis objects of the algebra in terms of an idempotent and an associated vector of basis functions. This has been shown for Clifford (2) and it can be extended to other algebras once the idempotent and the vector of basis functions have been identified. This has been done in many cases by Ablamowicz. This will now be developed to show how the inverse of a Clifford number can be obtained from the matrix representation.

4.3 Calculation of the inverse of a Clifford number

The matrix basis demonstrated above can be used to calculate the inverse of a Clifford number. In simple cases this can be used to obtain an algebraic formulation. For other cases the algebra is too complex to be clear, but the method can still be used to obtain the numerical value of the inverse. To apply the method it is necessary to know a basis matrix representation of the algebra being used.

The idea of the method is that the matrix representation will have a characteristic polynomial obeyed by the eigenvalues of the matrix and also by the matrix itself. There may also be a minimal polynomial which is a factor of the characteristic polynomial, which will have also be satisfied by the matrix. It is clear from the proceding section that if A is a matrix representation of c in a Clifford Algebra then if some function $f(A) = 0$ then the corresponding Clifford function $f(c) = 0$ must also be zero. In particular if $f(A) = 0$ is the characteristic or minimal polynomial of A, then $f(c) = 0$ implies that c also satisfies the same polynomial. Then if the inverse of the Clifford number, c^{-1} is to be found, then

$$c^{-1} f(c) = 0$$

provides a relationship for c^{-1} in terms of multiples a small number of low powers of c, with the maximum power one less than the order of the polynomial. The method suceeds unless the constant term in the polynomial is zero, which means that the inverse does not exist. For cases where the basis matrices are of order two, the inverse will be shown to be a linear function of c.

The method can be summed up as follows.

1. Find the matrix basis of the Clifford algebra.
2. Find the matrix representation of the Clifford number whose inverse is required.
3. Compute the characteristic or minimal polynomial.
4. Check for the existence of the inverse.
5. Compute the inverse using the coefficients from the polynomial.

Step 1 need only be done once for any Clifford algebra, and this can be done using the method in the previous section, where needed.

Step 2 is trivially a matter of accumulation of the correct multiples of the matrices.

Step 3 may involve the use of a computer algebra system to find the coefficients of the polynomial, if the matrix size is at all large.

Steps 4 and 5 are then easy once the coefficients are known.

The method will now be demonstrated using some examples.

4.3.1 Example 1: Clifford (2)

In this case the matrix basis for a member of the Clifford algebra

$$c = a_0 + a_1e_1 + a_2e_2 + a_3e_1e_2$$

was developed in the previous section as

$$A = \begin{pmatrix} a_0 + a_1 & a_2 + a_3 \\ a_2 - a_3 & a_0 - a_1 \end{pmatrix}$$

This matrix has the characteristic polynomial

$$X^2 - 2Xa_0 + a_0^2 - a_1^2 - a_2^2 + a_3^2 = 0$$

and therefore

$$X^{-1}(X^2 - 2Xa_0 + a_0^2 - a_1^2 - a_2^2 + a_3^2) = 0$$

and

$$X^{-1} = (2a_0 - X)/(a_0^2 - a_1^2 - a_2^2 + a_3^2) = 0$$

which provides a general solution to the inverse in this algebra.

$$c^{-1} = (2a_0 - c)/(a_0^2 - a_1^2 - a_2^2 + a_3^2) = 0$$

4.3.2 Example 2: Clifford (3)

A set of basis matrices for Clifford (3) as given by Abalmowicz and deduced are

$$\begin{array}{cc}
E_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & E_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & E_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} \\
E_1E_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & E_1E_3 = \begin{pmatrix} 0 & -j \\ -j & 0 \end{pmatrix} \\
E_2E_3 = \begin{pmatrix} j & 0 \\ 0 & -j \end{pmatrix} & E_1E_2E_3 = \begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix}
\end{array}$$

for the idempotent

$$f = \frac{(1 + e_1)}{2}, \text{ where } j^2 = -1.$$

The general member of the algebra

$$c_3 = a_0 + a_1e_1 + a_2e_2 + a_3e_3 + a_{12}e_1e_2 + a_{13}e_1e_3 + a_{23}e_2e_3 + a_{123}e_1e_2e_3$$

has the matrix representation

$$\begin{aligned}
A_3 &= a_0E_0 + a_1E_1 + a_2E_2 + a_3E_3 + a_{12}E_1E_2 \\
&\quad + a_{13}E_1E_3 + a_{23}E_2E_3 + a_{123}E_1E_2E_3 \\
&= \begin{pmatrix} a_0 + a_1 + ja_{23} + ja_{123} & a_2 - ja_3 + a_{12} - ja_{13} \\ a_2 + ja_3 - a_{12} - ja_{13} & a_0 - a_1 - ja_{23} + ja_{123} \end{pmatrix}
\end{aligned}$$

This has the characteristic polynomial

$$\begin{aligned}
& a_0^2 - a_1^2 - a_2^2 - a_3^2 + a_{12}^2 + a_{13}^2 + a_{23}^2 - a_{123}^2 \\
+ \quad & 2j(a_0 a_{123} - a_1 a_{23} - a_{12} a_3 + a_{13} a_2) \\
- \quad & 2(a_0 + j a_{123})X + X^2 = 0
\end{aligned}$$

and the expression for the inverse is

$$\begin{aligned}
X^{-1} = \; & (2a_0 + 2ja_{123} - X)/ \\
& (a_0^2 - a_1^2 - a_2^2 - a_3^2 + a_{12}^2 + a_{13}^2 + a_{23}^2 - a_{123}^2 \\
& +2j(a_0 a_{123} - a_1 a_{23} - a_{12} a_3 + a_{13} a_2))
\end{aligned}$$

Complex terms arise in two cases,

$$a_{123} \neq 0$$

and

$$(a_0 a_{123} - a_1 a_{23} - a_{12} a_3 + a_{13} a_2) \neq 0$$

Two simple cases have real minumum polynomials:

Zero and first grade terms only:

$$\begin{aligned}
A_1 &= a_0 E_0 + a_1 E_1 + a_2 E_2 + a_3 E_3 \\
&= \begin{pmatrix} a_0 + a_1 & a_2 - ja_3 \\ a_2 + ja_3 & a_0 - a_1 \end{pmatrix}
\end{aligned}$$

which has the minimum polynomial

$$a_0^2 - a_1^2 - a_2^2 - a_3^2 - 2a_0 X + X^2 = 0$$

which gives

$$X^{-1} = (2a_0 - X)/(a_0^2 - a_1^2 - a_2^2 - a_3^2)$$

Zero and second grade terms only (ie. the even subspace).

$$\begin{aligned}
A_2 = \; & a_0 E_0 + a_{12} E_1 E_2 + a_{13} E_1 E_3 + a_{23} E_2 E_3 \\
& \begin{pmatrix} a_0 + ja_{23} & a_{12} - ja_{13} \\ -a_{12} - ja_{13} & a_0 - ja_{23} \end{pmatrix}
\end{aligned}$$

which has minimum polynomial

$$a_0^2 + a_{23}^2 + a_{12}^2 + a_{13}^2 - 2a_0X + X^2 = 0$$

giving

$$X^{-1} = (2a_0 - X)/(a_0^2 + a_{23}^2 + a_{12}^2 + a_{13}^2)$$

This provides a general solution for the inverse together with two simple cases of wide usefulness.

4.3.3 Example 3: Clifford (2,2)

The following basis matrices are given by Ablamowicz[Ab98]

$$E_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$
$$E_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad E_4 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

for the idempotent

$$f = \frac{(1 + e_1e_3)(1 + e_1e_3)}{4}.$$

Note that this implies that the order of the basis members is such that e_1 and e_2 have square $+1$ and e_3 and e_4 have square -1. Other orderings are used by other authors. The remaining basis matrices can be deduced to be as follows.

Second Grade members

$$E_1E_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad E_1E_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
$$E_1E_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad E_2E_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$
$$E_2E_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad E_3E_4 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Third grade members

$$E_1E_2E_3 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad E_1E_2E_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

$$E_1E_3E_4 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad E_2E_3E_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Fourth grade member

$$E_1E_2E_3E_4 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Zero grade member (identity)

$$E_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The general member of the Clifford (2,2) algebra can be written as follows.

$$\begin{aligned} c_{22} \quad = \quad & a_0 + a_1e_1 + a_2e_2 + a_3e_3 + a_4e_4 + \\ & a_{12}e_1e_2 + a_{13}e_1e_3 + a_{14}e_1e_4 + a_{23}e_2e_3 + a_{24}e_2e_4 + a_{34}e_3e_4 \\ & + a_{123}e_1e_2e_3 + a_{124}e_1e_2e_4 + a_{134}e_1e_3e_4 + a_{234}e_2e_3e_4 + a_{1234}e_1e_2e_3e_4 \end{aligned}$$

This has the following matrix representation.

$$\begin{pmatrix} \begin{matrix} a_0 + a_{13} + \\ a_{24} - a_{1234} \end{matrix} & \begin{matrix} a_1 - a_3 + \\ a_{124} + a_{234} \end{matrix} & \begin{matrix} a_2 - a_4 - \\ a_{123} - a_{134} \end{matrix} & \begin{matrix} -a_{12} + a_{14} - \\ a_{23} - a_{34} \end{matrix} \\ \\ \begin{matrix} a_1 + a_3 + \\ a_{124} - a_{234} \end{matrix} & \begin{matrix} a_0 - a_{13} + \\ a_{24} + a_{1234} \end{matrix} & \begin{matrix} a_{12} - a_{14} - \\ a_{23} - a_{34} \end{matrix} & \begin{matrix} -a_2 + a_4 - \\ a_{123} - a_{134} \end{matrix} \\ \\ \begin{matrix} a_2 + a_4 - \\ a_{123} + a_{134} \end{matrix} & \begin{matrix} -a_{12} - a_{14} - \\ a_{23} + a_{34} \end{matrix} & \begin{matrix} a_0 + a_{13} - \\ a_{24} + a_{1234} \end{matrix} & \begin{matrix} a_1 - a_3 - \\ a_{124} - a_{234} \end{matrix} \\ \\ \begin{matrix} a_{12} + a_{14} - \\ a_{23} + a_{34} \end{matrix} & \begin{matrix} -a_2 - a_4 - \\ a_{123} + a_{134} \end{matrix} & \begin{matrix} a_1 + a_3 - \\ a_{124} + a_{234} \end{matrix} & \begin{matrix} a_0 - a_{13} - \\ a_{24} - a_{1234} \end{matrix} \end{pmatrix}$$

In this case it is possible to generate the characteristic equation using computer algebra. However, it is too complex to be of practical use. Instead here are numerical examples of the use of the method to calculate the inverse. For the case where

$$n1 = 1 + e_1 + e_2 + e_3 + e_4$$

then the matrix representation is

$$N_1 = E_0 + E_1 + E_2 + E_3 + E_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & -2 & 2 & 1 \end{pmatrix}$$

This has the minimum polynomial

$$X^2 - 2X + 1 = 0$$

so that

$$X^{-1} = 2 - X$$

and

$$n_1^{-1} = 2 - n_1 = 1 - e_1 - e_2 - e_3 - e_4$$

For

$$n_2 = 1 + e_1 + e_2 + e_3 + e_4 + e_1 e_2$$

the matrix representation is

$$N_2 = I + E_1 + E_2 + E_3 + E_4 + E_1 E_2 = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 2 & 1 & 1 & 0 \\ 2 & -1 & 1 & 0 \\ 1 & -2 & 2 & 1 \end{pmatrix}$$

This has the minimum polynomial

$$X^4 - 4X^3 + 8X^2 - 8X - 4 = 0$$

so that

$$X^{-1} = \frac{X^3 - 4X^2 + 8X - 8}{4}$$

and

$$n_2^{-1} = \frac{n_2^3 - 4n_2^2 + 8n_2 - 8}{4}$$

This expression can be evaluated easily using a computer algebra system for Clifford algebra such as described in Fletcher[Fl01]. The result is

$$\begin{aligned} n_2^{-1} &= -0.5 + 0.5e_1 + 0.5e_2 - 0.5e_1e_2 - 0.5e_1e_3 \\ &\quad -0.5e_1e_4 + 0.5e_2e_3 + 0.5e_2e_4 - 0.5e_1e_2e_3 - 0.5e_1e_2e_4 \end{aligned}$$

Note that in some cases the inverse is linear in the original Clifford number, and in others it is nonlinear.

4.3.4 Conclusion

The paper has demonstrated a method for the calculation of inverses of Clifford numbers by means of the matrix representation of the corresponding Clifford algebra. The method depends upon the calculation of the basis matrices for the algebra. This can be done from an idempotent for the algebra if the matrices are not already available. The method provides an easy check on the existence of the inverse. For simple systems a general algebraic solution can be found and for more complex systems the algebra of the inverse can be generated and evaluated numerically for a particular example, given a system of computer algebra for Clifford algebra.

Chapter 5

Package for Algebraic Function Fields

PAFF is a Package for Algebraic Function Fields in one variable by Gaétan Haché

PAFF is a package written in Axiom and one of its many purpose is to construct geometric Goppa codes (also called algebraic geometric codes or AG-codes). This package was written as part of Gaétan's doctorate thesis on "Effective construction of geometric codes": this thesis was done at Inria in Rocquencourt at project CODES and under the direction of Dominique LeBrigand at Universit Pierre et Marie Curie (Paris 6). Here is a résumé of the thesis.

It is well known that the most difficult part in constructing AG-code is the computation of a basis of the vector space "L(D)" where D is a divisor of the function field of an irreducible curve. To compute such a basis, PAFF used the Brill-Noether algorithm which was generalized to any plane curve by D. LeBrigand and J.J. Risler (see [LR88]). In [Ha96] you will find more details about the algorithmic aspect of the Brill-Noether algorithm. Also, if you prefer, as I do, a strictly algebraic approach, see [Ha95]. This is the approach I used in my thesis ([Ha96]) and of course this is where you will find complete details about the implementation of the algorithm. The algebraic approach use the theory of algebraic function field in one variable : you will find in [St93] a very good introduction to this theory and AG-codes.

It is important to notice that PAFF can be used for most computation related to the function field of an irreducible plane curve. For example, you can compute the genus, find all places above all the singular points, compute the adjunction divisor and of course compute a basis of the vector space L(D) for any divisor D of the function field of the curve.

There is also the package PAFFFF which is especially designed to be used over finite fields. This package is essentially the same as PAFF, except that the computation are done over “dynamic extensions” of the ground field. For this, I used a simplify version of the notion of dynamic algebraic closure as proposed by D. Duval (see [Du95]).

Example 1

This example compute the genus of the projective plane curve defined by:

```
  5    2 3      4
 X  + Y Z  + Y Z  = 0
```

over the field GF(2).

First we define the field GF(2).

```
K:=PF 2
R:=DMP([X,Y,Z],K)
P:=PAFF(K,[X,Y,Z],BLQT)
```

We defined the polynomial of the curve.

```
C:R:=X**5 + Y**2*Z**3+Y*Z**4
```

We give it to the package PAFF(K,[X,Y,Z]) which was assigned to the variable P.

```
setCurve(C)$P
```

Chapter 6

Groebner Basis

Groebner Basis

Chapter 7

Greatest Common Divisor

Greatest Common Divisor

Chapter 8

Polynomial Factorization

Polynomial Factorization

Chapter 9

Cylindrical Algebraic Decomposition

Cylindrical Algebraic Decomposition

Chapter 10

Pade approximant

Pade approximant

Chapter 11

Schwartz-Zippel lemma and testing polynomial identities

Schwartz-Zippel lemma and testing polynomial identities

Chapter 12

Chinese Remainder Theorem

Chinese Remainder Theorem

Chapter 13

Gaussian Elimination

Gaussian Elimination

Chapter 14

Diophantine Equations

Diophantine Equations

Bibliography

[Ab98] Ablamowicz Rafal, "Spinor Representations of Clifford Algebras: A Symbolic Approach", Computer Physics Communications Vol. 115, No. 2-3, December 11, 1998, pages 510-535.

[Alt05] Altmann, Simon L. Rotations, Quaternions, and Double Groups Dover Publications, Inc. 2005 ISBN 0-486-44518-6

[Ber95] Laurent Bertrand. Computing a hyperelliptic integral using arithmetic in the jacobian of the curve. *Applicable Algebra in Engineering, Communication and Computing*, 6:275-298, 1995

[Bro90] M. Bronstein. "On the integration of elementary functions" *Journal of Symbolic Computation* 9(2):117-173, February 1990

[Bro91] M. Bronstein. The Risch differential equation on an algebraic curve. In S.Watt, editor, *Proceedings of ISSAC'91*, pages 241-246, ACM Press, 1991.

[Bro97] M. Bronstein. *Symbolic Integration I–Transcendental Functions.* Springer, Heidelberg, 1997

[Br98] Bronstein, Manuel "Symbolic Integration Tutorial" INRIA Sophia Antipolis ISSAC 1998 Rostock

[Bro98] M. Bronstein. The lazy hermite reduction. Rapport de Recherche RR-3562, INRIA, 1998

[CS03] Conway, John H. and Smith, Derek, A., "On Quaternions and Octonions" A.K Peters, Natick, MA. (2003) ISBN 1-56881-134-9

[Dal03] Daly, Timothy, "The Axiom Wiki Website" `http://axiom.axiom-developer.org`

[Dal09] Daly, Timothy, "The Axiom Literate Documentation" `http://axiom.axiom-developer.org/axiom-website/documentation.html`

[Du95] Duval, D. "Evaluation dynamique et clôture algébrique en Axiom". Journal of Pure and Applied Algebra, no99, 1995, pp. 267–295.

[Fl01] Fletcher, John P. "Symbolic processing of Clifford Numbers in C++", Paper 25, AGACSE 2001.

[Fl09] Fletcher, John P. "Clifford Numbers and their inverses calculated using the matrix representation." Chemical Engineering and Applied Chemistry, School of Engineering and Applied Science, Aston University, Aston Triangle, Birmingham B4 7 ET, U. K. `www.ceac.aston.ac.uk/research/staff/jpf/papers/paper24/index.php`

[Ga95] Garcia, A. and Stichtenoth, H. "A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound" Invent. Math., vol. 121, 1995, pp. 211–222.

[Ha1896] Hathway, Arthur S., "A Primer Of Quaternions" (1896)

[Ha95] Haché, G. "Computation in algebraic function fields for effective construction of algebraic-geometric codes" Lecture Notes in Computer Science, vol. 948, 1995, pp. 262–278.

[Ha96] Haché, G. "Construction effective des codes géométriques" Thése de doctorat de l'Université Pierre et Marie Curie (Paris 6), Septembre 1996.

[Her1872] E. Hermite. Sur l'intégration des fractions rationelles. *Nouvelles Annales de Mathématiques* (2^{eme} série), 11:145-148, 1872

[HI96] Huang, M.D. and Ierardi, D. "Efficient algorithms for Riemann-Roch problem and for addition in the jacobian of a curve" Proceedings 32nd Annual Symposium on Foundations of Computer Sciences. IEEE Comput. Soc. Press, pp. 678–687.

[HL95] Haché, G. and Le Brigand, D. "Effective construction of algebraic geometry codes" IEEE Transaction on Information Theory, vol. 41, n27 6, November 1995, pp. 1615–1628.

[JS92] Richard D. Jenks and Robert S. Sutor "AXIOM: The Scientific Computation System" Springer-Verlag, Berlin, Germany / Heildelberg, Germany / London, UK / etc., 1992 ISBN 0-387-97855-0 (New York), 3-540-97855-0 (Berlin) 742pp LCCN QA76.95.J46 1992

[Knu84] Knuth, Donald, *The TEXbook*
Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1984. ISBN 0-201-13448-9

[Kn92] Knuth, Donald E., "Literate Programming" Center for the Study of Language and Information ISBN 0-937073-81-4 Stanford CA (1992)

[La86] Lamport, Leslie, *LaTeX: A Document Preparation System,*
Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1986. ISBN 0-201-15790-X

[LR88] Le Brigand, D. and Risler, J.J. "Algorithme de Brill-Noether et codes de Goppa" Bull. Soc. Math. France, vol. 116, 1988, pp. 231–253.

[LR90] Daniel Lazard and Renaud Rioboo. Integration of rational functions: Rational coputation of the logarithmic part *Journal of Symbolic Computation*, 9:113-116:1990

[Lio1833a] Joseph Liouville. Premier mémoire sur la détermination des intégrales dont la valeur est algébrique. *Journal de l'Ecole Polytechnique*, 14:124-148, 1833

[Lio1833b] Joseph Liouville. Second mémoire sur la détermination des intégrales dont la valeur est algébrique. *Journal de l'Ecole Polytechnique*, 14:149-193, 1833

[Mul97] Thom Mulders. A note on subresultants and a correction to the lazard/rioboo/trager formula in rational function integration *Journal of Symbolic Computation*, 24(1):45-50, 1997

[Ost1845] M.W. Ostrogradsky. De l'intégration des fractions rationelles. *Bulletin de la Classe Physico-Mathématiques de l'Acaeémie Impériale des Sciences de St. Pétersbourg,* IV:145-167,286-300, 1845

[Pu09] Puffinware LLC "Singular Value Decomposition (SVD) Tutorial" `www.puffinwarellc.com/p3a.htm`

[Ra03] Ramsey, Norman "Noweb – A Simple, Extensible Tool for Literate Programming" `www.eecs.harvard.edu/~nr/noweb`

[Ris68] Robert Risch. On the integration of elementary functions which are built up using algebraic operations. Research Report SP-2801/002/00, System Development Corporation, Santa Monica, CA, USA, 1968

[Ris69a] Robert Risch. "Further results on elementary functions" Research Report RC-2042, IBM Research, Yorktown Heights, NY, USA, 1969

[Ris69b] Robert Risch, "The problem of integration in finite terms" *Transactions of the American Mathematical Society* 139:167-189, 1969

[Ris70] Robert Risch. "The solution of problem of integration in finite terms" *Transactions of the American Mathematical Society* 76:605-608, 1970

[Ris79] Robert Risch. "Algebraic properties of the elementary functions of analysis" *American Journal of Mathematics*, 101:743-759, 1979

[Ro72] Maxwell Rosenlicht. Integration in finite terms. *American Mathematical Monthly*, 79:963-972, 1972

[Ro77] Michael Rothstein. "A new algorithm for the integration of exponential and logarithmic functions" In *Proceedings of the 1977 MACSYMA Users Conference*, pages 263-274. NASA Pub CP-2012, 1977

[St93] Stichtenoth, H. "Algebraic function fields and codes" Springer-Verlag, 1993, University Text.

[Ta1890] Tait, P.G.,*An Elementary Treatise on Quaternions* C.J. Clay and Sons, Cambridge University Press Warehouse, Ave Maria Lane 1890

[Tr76] Trager, Barry "Algebraic factoring and rational function integration" In Proceedings of SYMSAC'76 pages 219-226, 1976

[Tr84] Trager Barry *On the integration of algebraic functions*, PhD thesis, MIT, Computer Science, 1984

[vH94] M. van Hoeij. "An algorithm for computing an integral basis in an algebraic function field" *J. Symbolic Computation* 18(4):353-364, October 1994

[Wa03] Watt, Stephen, "Aldor", `www.aldor.org`

[We71] André Weil, *Courbes algébriques et variétés Abeliennes* Hermann, Paris, 1971

[Wo09] Wolfram Research, `mathworld.wolfram.com/Quaternion.html`

[Yu76] D.Y.Y. Yun. "On square-free decomposition algorithms" *Proceedings of SYMSAC'76* pages 26-35, 1976

Chapter 15

Index