INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460392 - I

SECOND SCHOOL ON ADVANCED TECHNIQUES
IN COMPUTATIONAL PHYSICS
(18 January - 12 February 1988)

SMR.282/23

Algebraic and Symbolic Methods

A.C.HEARN

The RAND Corporation, Santa Monica, USA

1

ALGEBRAIC AND SYMBOLIC METHODS

Anthony C. Hearn
The RAND Corporation

2

ALGEBRAIC MANIPULATION PROVIDES FOR

- Exact results
- Removal of instabilities
- Preparation and optimization of numeric codes
- ''Symbol crunching''

3

CHARACTERISTICS OF COMPUTER ALGEBRA

- Algebra rather than analysis
- Constructive rather than existential algebra
- Simplification plays a key role

4

IMPORTANT AREAS OF CONCERN

- Can you get a result at all?
- Can you develop algorithms or heuristics to do it?
- Can you get it in reasonable time and space?
- Is result comprehensible?

```
      t1 = (((24*(2*((-y8+y7-y6+y5-y4+y3-y2+y1)*(z8+z7-z6-z5-z4-z3+z2+z1
     1  )+(-y8-y7+y6+y5-y4-y3+y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)-(y8+y7-y
     2  6-y5-y4-y3+y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)-(y8-y7+y6-y5-y4+y3
     3  -y2+y1)*(-z8-z7+z6+z5-z4-z3+z2+z1))-2*(y8-y7+y6-y5-y4+y3-y2+y1)
     4  *(z8+z7-z6-z5-z4-z3+z2+z1)+2*(y8+y7-y6-y5-y4-y3+y2+y1)*(z8-z7+z
     5  6-z5-z4+z3-z2+z1)-2*(-y8+y7-y6+y5-y4+y3-y2+y1)*(z8-z7+z6-z5-z4+
     6  z3-z2+z1)-2*(-y8-y7+y6+y5-y4-y3+y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1
     7  )+2*(y8-y7+y6-y5-y4+y3-y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)+2*(y8-
     8  y7+y6-y5-y4+y3-y2+y1)*(-z8-z7+z6+z5-z4-z3+z2+z1))**2+24*(2*(2*(
     9  -y8+y7-y6+y5-y4+y3-y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)-2*(y8-y7+y6
     :  -y5-y4+y3-y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1))+2*(2*(y8-y7+y6-y5-
     ;  y4+y3-y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)-2*(-y8+y7-y6+y5-y4+y3-y
     <  2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)))*(2*((-y8-y7+y6+y5-y4-y3+y2+y1
     =  )*(z8+z7-z6-z5-z4-z3+z2+z1)-(y8+y7-y6-y5-y4-y3+y2+y1)*(-z8-z7+z
     >  6+z5-z4-z3+z2+z1))-2*(y8-y7+y6-y5-y4+y3-y2+y1)*(z8+z7-z6-z5-z4-
     ?  z3+z2+z1)+2*(y8+y7-y6-y5-y4-y3+y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)
     @  -2*(-y8-y7+y6+y5-y4-y3+y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)+2*(y8-y
     1  7+y6-y5-y4+y3-y2+y1)*(-z8-z7+z6+z5-z4-z3+z2+z1)))/3+2*(12*(2*((
     2  -y8+y7-y6+y5-y4+y3-y2+y1)*(z8+z7-z6-z5-z4-z3+z2+z1)+(-y8-y7+y6+
     3  y5-y4-y3+y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)-(y8+y7-y6-y5-y4-y3+y2
     4  +y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)-(y8-y7+y6-y5-y4+y3-y2+y1)*(-z8-
     5  z7+z6+z5-z4-z3+z2+z1))-2*(-y8+y7-y6+y5-y4+y3-y2+y1)*(z8+z7-z6-z
     6  5-z4-z3+z2+z1)-2*(-y8+y7-y6+y5-y4+y3-y2+y1)*(z8-z7+z6-z5-z4+z3-
     7  z2+z1)+2*(y8+y7-y6-y5-y4-y3+y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)+2
     8  *(y8-y7+y6-y5-y4+y3-y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)-2*(-y8-y7
     9  +y6+y5-y4-y3+y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)+2*(-y8+y7-y6+y5-
     :  y4+y3-y2+y1)*(-z8-z7+z6+z5-z4-z3+z2+z1))**2+12*(2*(2*(-y8+y7-y6
     ;  +y5-y4+y3-y2+y1)*(z8-z7+z6-z5-z4+z3-z2+z1)-2*(y8-y7+y6-y5-y4+y3
     <  -y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1))+2*(2*(y8-y7+y6-y5-y4+y3-y2+
     =  y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)-2*(-y8+y7-y6+y5-y4+y3-y2+y1)*(z8
     >  -z7+z6-z5-z4+z3-z2+z1)))*(2*((-y8-y7+y6+y5-y4-y3+y2+y1)*(z8+z7-
     ?  z6-z5-z4-z3+z2+z1)-(y8+y7-y6-y5-y4-y3+y2+y1)*(-z8-z7+z6+z5-z4-z
     @  3+z2+z1))-2*(-y8+y7-y6+y5-y4+y3-y2+y1)*(z8+z7-z6-z5-z4-z3+z2+z1
     1  )+2*(y8+y7-y6-y5-y4-y3+y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)-2*(-y8
     2  -y7+y6+y5-y4-y3+y2+y1)*(-z8+z7-z6+z5-z4+z3-z2+z1)+2*(-y8+y7-y6+
     3  y5-y4+y3-y2+y1)*(-z8-z7+z6+z5-z4-z3+z2+z1))))/60
```

```
t1 =(128.*(3.*(y8*z7-y8*z6+y8*z3-y8*z2-y7*z8+y7*z5-y7*z3+y7
.     *z2+y6*z8-y6*z5+y6*z3-y6*z2-y5*z7+y5*z6-y5*z3+y5*z2-y3*z8
.     +y3*z7-y3*z6+y3*z5+y2*z8-y2*z7+y2*z6-y2*z5)**2+(y8*z7-y8*
.     z6-y8*z3+y8*z2-y7*z8+y7*z5+y7*z3-y7*z2+y6*z8-y6*z5-y6*z3+
.     y6*z2-y5*z7+y5*z6+y5*z3-y5*z2+y3*z8-y3*z7+y3*z6-y3*z5-y2*
.     z8+y2*z7-y2*z6+y2*z5)**2))/15.
```

7

BRIEF HISTORY

- 1960's:
  - programs and systems emerge
  - simple algorithms
  - limited facilities
  - limited user base
- 1970's:
  - more sophisticated algorithms
  - systems mature
  - user base grows
- (late) 1980's:
  - small system implementations widely available
  - algorithms continue to improve
  - user base very large

8

ALGEBRAIC SYSTEM CAPABILITIES

- Integer arithmetic
- Real arithmetic
- Polynomial arithmetic
- Rational function arithmetic
- Power series manipulation
- Matrix manipulation
- Dirac trace algebra
- Differentiation
- Integration

ALGEBRAIC SYSTEM CAPABILITIES (Continued)

- Linear Equations
- Non-linear equations
- Differential equations
- Integral equations
- Special functions
  - * Pattern matching
  - * User tools

ALGEBRAIC SYSTEM APPLICATIONS

* Quantum electrodynamics
+ Celestial mechanics
+ General relativity
- Plasma physics
- Fluid mechanics
- Statistical mechanics
- Structural mechanics
- Quantum chemistry
- Electron optics

11

ALGEBRAIC SYSTEM APPLICATIONS (Continued)

- Numerical analysis
- Program transformations
- Pure mathematics

12

METHODS FOR WRITING ALGEBRA SYSTEMS

- Written in assembly language
- Written in a high-level language (BCPL, C, FORTRAN, LISP ...)
- Written in own language (but embedded in another language)
- Interpreters rather than compilers
- Small computer implementations differ only in scale (but this difference is rapidly disappearing)

AVAILABLE COMPUTER ALGEBRA SYSTEMS

- ALTRAN (based on ANSI FORTRAN):
  W. S. Brown, Computing Information Library, Bell Telephone Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974

- CAMAL (host language BCPL):
  J. P. Fitch, School of Mathematics, University of Bath, Claverton Down, Bath BA2 7AY, England.

- FORMAC (based on FORTRAN or PL/I):
  K. A. Bahr, Gesellschaft fuer Mathematik und Datenverarbeitung G/IFV, Rheinstrasse 75, D-6100 Darmstadt, Federal Republic of Germany.

- MACSYMA (host language LISP):
  Symbolics, Inc., 257 Vassar Street, Cambridge MA 02139.

- MAPLE (based on C):
  K. O. Geddes, Computer Science Department, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

13

AVAILABLE COMPUTER ALGEBRA SYSTEMS (Continued)

- muMATH (Intel 8080, 8086 and Zilog Z-80 based):
  The Soft Warehouse, P.O. Box 11174, Honolulu, HI 96828.

- REDUCE (bootstrapped from a LISP kernel):
  A. C. Hearn, The RAND Corporation, 1700 Main Street, Santa Monica, CA 90406.

- SAC-2 (based on ANSI FORTRAN):
  G. E. Collins, Computer Science Department, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706.

- SCRATCHPAD (based on LISP):
  R. D. Jenks, IBM T. J. Watson Research Center, P. O. Box 218, Yorktown Heights, NY 10598.

14

ELEMENTARY DATA STRUCTURES

- Integers
- Reals
- Polynomials
- Rational functions
- Power series
- General expressions

REPRESENTATION OF INTEGERS

- Fixed radix : $a = \sum_{i=0}^{n-1} a_i x^i$, $a_{n-1} \neq 0$

  Implement as (reversed) list

  $(a_0, \ldots, a_{n-1})$

  Addition : pairwise addition of digits (with carry)

- Modular : $a \Leftrightarrow (v_0, \ldots, v_{n-1})$

  wrt $(p_0, \ldots, p_{n-1})$

  where p_i are relatively prime.

  a is unique if $0 \leq a < \prod_{i=0}^{n-1} p_i$

- P-adic : $a = \sum_{i=0}^{n-1} u_i p^i$, p prime

UNIVARIATE POLYNOMIALS

$$x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$$

$$\oplus$$

$$3x^6 + 5x^4 \qquad - 4x^2 - 9x + 21$$

=>

- dense representation

-5, 2, 8, -3, -3, 0, 1, 0, 1

21, -9, -4, 0, 5, 0, 3

16, -7, 4 -3, 2, 0, 4, 0, 1

- sparse representation

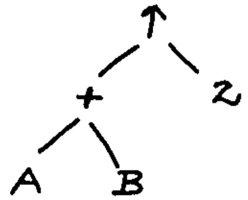
(1, 8), (1, 6), (-3, 4), (-3, 3),

(8, 2), (2, 1), (-5, 0)

MULTIVARIATE POLYNOMIALS

- Trees

$(A+B)^2$ =>



- Distributive representation

$$u(x_1, \ldots, x_n) = \sum_{\nu=0}^{m} u_\nu x_1^{\nu_1} \ldots x_n^{\nu_n}$$

- Recursive representation

$$u(x_1, \ldots, x_n) = \sum_{\nu=0}^{m} u_\nu(x_2, \ldots, x_n)\, x_1^{\nu}$$

$$u(x) = \sum_{\gamma=0}^{n} u_\gamma x^\gamma$$

19

$$u(x) = \prod_i \left(\sum_{\gamma_i=0}^{n_i} u_{\gamma_i} x^{\gamma_i} \right)^{\alpha_i}$$

CATEGORIES OF STRUCTURING

20

- Structure preserving
- Structure determining
- Structure reducing
- Structure displaying

21

INTEGRATION METHODS

- Pattern matching
- Heuristic programs
- Special purpose algorithms
- General algorithms (Risch)

22

INTEGRATION BY PATTERN MATCHING

$(x \uparrow k \; e \uparrow -x)$

```
Operator igl;

Linear igl;

Let igl(1,x) = x, igl(x,x) = x**2/2;

For all k let igl(x**k,x)
   = x**(k+1)/(k+1);

Let igl(e**(-x),x)= -e**(-x),
     igl(x*e**(-x),x) =
        -x*e**(-x)-e**(-x);

For all k let igl(x**k*e**(-x),x) =
   -x**k*e**(-x)
   +k*igl(x**(k-1)*e**(-x),x);
```

ALGORITHMIC INTEGRATION

- Rational case :

$$\int f(x)\,dx = q(x) + \sum_i \int \frac{a_i(x)}{b_i(x)}\,dx,$$

where b_i have roots of multiplicity 1

i.e.,

$$\int f(x)\,dx = q(x) + \sum_i c_i \log(r_i(x)),$$

where $q(x)$, $r_i(x)$ are rational and c_i are constants.

(Liouville)

RATIONAL FUNCTION INTEGRATION

1) $Q(x)/S(x) = P(x) + R(x)/S(x)$, $\deg(R) < \deg(S)$.

2) $S(x) = S_1(x)\,S_2^2(x)\ldots S_k^k$, $\gcd(S_i, S_j) = 1$, $i \neq j$.

3) $\exists$ A, B such that

$$A S_k^k + B S_1 S_2^2 \ldots S_{k-1}^{k-1} = 1.$$

4) Dividing by S and multiplying by R :

$$\frac{AR}{S_1 S_2^2 \ldots S_{k-1}^{k-1}} + \frac{BR}{S_k^k} = \frac{R}{S}$$

$$\Rightarrow \frac{R}{S} = \frac{A_1(x)}{S_1} + \frac{A_2(x)}{S_2^2} + \cdots + \frac{A_k(x)}{S_k^k}.$$

Reduction of $\int \frac{A_i}{S_i^i}$:

$\gcd(S_i, S_i') = 1$

$\Rightarrow \exists\ B, C$ such that

$B S_i + C S_i' = 1$.

So $\frac{B A_i}{S_i^{i-1}} + \frac{C A_i S_i'}{S_i^i} = \frac{A_i}{S_i^i}$.

$$\text{So } \int \frac{A_i}{S_i^i}\,dx = \int \frac{D}{S_i^{i-1}}\,dx + \int \frac{E S_i'}{S_i^i}\,dx$$

$$= -\frac{E}{(i-1)S_i^{i-1}} + \int \frac{D + E_i'/(i-1)}{S_i^{i-1}}.$$

$$\int \frac{R(x)}{S(x)}\,dx = \frac{T(x)}{U(x)} + \sum_{j=1}^{k} \frac{V_j(x)}{W_j(x)}\,dx$$

$$= q(x) + \sum_i c_i \log(r_i(x)).$$

- General case : (Risch, 1968)

Extension of Liouville result

Uses algebraic independence

Given $f(x) \in R(x, \log(x), e^x, \ldots)$, what can $\int f(x)\,dx$ be ?

Answer: if it exists in closed terms, must be of form:

$$\int f(x)\,dx = q(x) + \sum_i c_i \log(r_i(x)),$$

where $q(x), r_i(x) \in R$ and c_i are constants.

RISCH - NORMAN ALGORITHM
(1976)

1) Start with integrand as a rational form in a set of independent objects $x, e^x, \log(x)$, and so on.

e.g. consider $\int \log(x)\,dx = \int y\,dx$.
where $y = \log(x)$.

2) Substitute a purely formal sum for integral.
i.e. if integral exists, it is of form

$$u(x,y) = \sum u_{ij}\, x^i y^j .$$

3) Differentiate integral and equate result to integrand

$$\log(x) = u'(x,y)$$
$$= \sum u_{ij}\,(i\,x^{i-1} y^j + j\,x^{i-1} y^{j-1}).$$

4) Manipulate sum to obtain recurrence relation for coefficients

$$\log(x) = \sum x^i y^j \{ (i+1)\, u_{i+1,j} + (j+1)\, u_{i+1,j+1} \}$$

5) Use recurrence relation to take out term of highest degree

$$i=0,\ j=1 : \quad 1 = u_{1,1} + \cancel{2u_{1,2}}$$

6) Use 5) to solve residual eqn:

=>

$$\int \log(x)\,dx = x\log(x) - \int dx$$
$$= x\log(x) - x .$$

Now consider $\int x^2 e^{x^2} dx$.

Integral, if it exists, is of form

$$u(x,y) = \sum u_{ij} x^i y^j,$$

where $y = e^{x^2}$.

$$x^2 e^{x^2} = u'(x,y) = \sum u_{ij}\left(i x^{i-1} y^j + 2j x^{i+1} y^j\right)$$

$$= \sum x^i y^j \left\{ (i+1)\, u_{i+1,j} + 2j\, u_{i-1,j} \right\}.$$

$i=2,\ j=1$: $\quad 1 = 3\cancel{u_{3,1}} + 2u_{1,1} \Rightarrow u_{1,1} = \frac{1}{2}$

$$\int x^2 e^{x^2} dx = \tfrac{1}{2} x e^{x^2} - \tfrac{1}{2}\int e^{x^2} dx \qquad *$$

$i=0,\ j=1$: $\quad -\frac{1}{2} = u_{1,1} + 2\cancel{u_{-1,1}}$

inconsistency

$$\int \frac{\sqrt{x + \sqrt{x^2 + a^2}}}{x}\, dx$$

$$\int \frac{\sqrt{x+\sqrt{x^2+a^2}}}{x}\,dx$$

$$= \log(-yz\sqrt{a} + ya - za\sqrt{a} + zx\sqrt{a} + a^2)$$

$$+ \log(-yzi\sqrt{a} - ya + zai\sqrt{a} + zxi\sqrt{a} + a^2)$$

$$- \quad a(1+i)\log x + 2z$$

$$\text{where } y = \sqrt{x^2+a^2}$$

$$z = \sqrt{x+y}$$

31

INTEGRATION REFERENCES

G. H. Hardy ‘‘The Integration of Functions of a Single Variable’’ (CUP 1916)

J. Moses ‘‘Symbolic Integration: The Stormy Decade’’ CACM 8 (1971) 548-560

R. Risch ‘‘The Problem of Integration in Finite Terms’’ Trans. AMS 139 (1969) 167-189

A. C. Norman and P. M. A. Moore ‘‘Implementing the New Risch Algorithm’’, Proc. 4th Intern. Colloq. on Advanced Computing Meth. in Theor. Phys., CNRS, Marseilles (1977)

J. H. Davenport, ‘‘On the Integration of Algebraic Functions’’, Lecture Notes in Comp. Science 102, Springer Verlag, Berlin, 1981.

A. C. Norman and J. H. Davenport, ‘‘Symbolic Integration - The Dust Settles?’’, Proc. EUROSAM ’79, published as Lecture Notes in Comp. Science 72, Springer Verlag, Berlin (1981) 398-407

32

INFRASTRUCTURE OF THE (EARLY) 1970's

33

- Centralized computing equipment
- Centralized job entry
- Difficult access
- Centralized printing
- Batch processing
- Limited storage capacity

INFRASTRUCTURE OF THE (LATE) 1980's

34

- Local workstations
- Communications network
- Local and remote printing
- Remote large scale processors
- Easy access
- Interactive computing
- Adequate storage capacity

MODERN ALGEBRA SYSTEMS ARE WELL-MATCHED TO PERSONAL COMPUTERS

- Immediate access
- Integrated environment
- Interactive step-wise command development
- Graphics an important adjunct

but ...

- Need large address spaces
- Need large memories

35

WORKSTATION CAPABILITIES

- Voice processing
- Text processing
- Mail
- Numerical processing
- Graphical processing
- Symbolic processing

36

THE FUTURE

- Wide Availability
- Extensive facilities
- Integrated Environments
- Effective hybrid calculations
- "Electronic Bateman"

References

Buchberger, B. et al, eds, "Computer Algebra - Symbolic and Algebraic Computation", Computing - Computing Supplementum 4, Springer-Verlag, Vienna, 1982.

Davenport, J.H., E. Siret & E. Tournier, "Computer Algebra". Academic Press, 1988. (An earlier edition of this is also available in French as "Calcul Formel", Masson 1986.)

Yun, D. Y. Y., and Stoutemyer, D. R., "Symbolic Mathematical Computation" in Encyclopedia of Computer Science and Technology (Belzer, J. et al, eds), Vol 15, 235-310 (1980).

SIGSAM Bulletin, quarterly publication of the ACM, New York.

Proceedings of various SIGSAM conferences: SYMSAM 66, SYMSAM 71, EUROSAM 74, SYMSAC 76, EUROSAM 79, SYMSAC 81, EUROCAM 82, EUROCAL 83, EUROSAM 84, EUROCAL 85, SYMSAC 86. The last five "euro" conferences were published by Springer-Verlag, Heidelberg, the remainder by ACM, New York.

The Second RIKEN International Symposium on Symbolic and Algebraic Computation by Computers, N. Inada and T. Soma, Editors, World Scientific, Singapore.

Journal of Symbolic Computation, B. Buchberger, Editor, published quarterly by Academic Press.

Occasional articles on algebraic computation (mainly applications) may be found in:

Computer Physics Communications, monthly - bimonthly publication of North-Holland, Amsterdam.

Journal of Computational Physics, monthly publication of Academic Press, New York.

Trans. on Math. Software, quarterly publication of the ACM, New York.

Books on specific systems are also available, for example:

R.H. Rand, "Computer Algebra in Applied Mathematics: An Introduction to MACSYMA", Pitman, 1984.

Gerhard Rayna, "REDUCE Software for Algebraic Computation", Springer-Verlag, 1987.

RECENT REVIEWS OF ALGEBRAIC COMPUTATION APPLICATIONS

Calmet, J., J.A. Van Hulzen, "Computer Algebra Applications", Computing - Computing Supplementum 4, Springer-Verlag, Vienna, 1982.

Davenport, J.H., "Mathematics for Problem Solving Environments: Symbolic Computation", North Holland (in press).

Fitch, J.P., "Solving Algebraic Problems with REDUCE", J. Symb. Comp. 1 (1985) 211-227.