Regular Array Expansions in Null Arrays, with Applications:

Zero Is Not Nothing

by

Ronald I. Frank

**Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Professional Studies
in Computing**

at

School of Computer Science and Information Systems

Pace University

June 2002

UMI Number: 3064835

UMI®

UMI Microform 3064835

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

We hereby certify that this dissertation, submitted by Ronald I. Frank, satisfies the dissertation requirements for the degree of Doctor of Professional Studies and has been approved.

Dr. Fred Grossman **Date**

Chairperson of Dissertation Committee

Dr. Michael Gargano **Date**

Dissertation Committee Member

Dr. Allen Stix **Date**

Dissertation Committee Member

School of Computer Science and Information Systems

Pace University 2002

Abstract

I show how a simple observation about mapping the binomial expansion to regular arrays leads to a decomposition of N-D equilateral regular arrays (ERAs) of axis length n, in terms of ERAs of all dimensions from 0 to N but of axis length one less (n-1) or one more (n+1). I then recurse the decompositions and derive a single canonical expansion of an N-D ERA of axis length n, in terms of null arrays.

I recapitulate these arguments and generalize them for non-equilateral regular arrays (NERAs). I call the final canonical decomposition of a NERA in terms of only null arrays the "master equation" because all other results can be derived from it. I derive many examples and subsidiary results along the way including the expansions in nulls for an arbitrary inner product and for a catenation of two conformable arrays. I also have found explanations for some open array questions.

There are four chapters on applications: one suggests a new operation on arrays – array division, and one suggests analyzing hyper cube machines and generalized hyper cube machines using these methods. The third applies these methods to OLAP hierarchical data, and the last one suggests enlarging computer languages to properly allow null arrays.

This is a rich set of results, so I indicate a number of possible new avenues of research applying these results. This includes array valued polynomials, fractional and complex length arrays, "ragged" arrays, complex and fractional dimension arrays, and GHC connectivity .

Key Words

Null Array, Binomial Array Expansion, Processor Connectivity, N Dimensional Machines, Null Data Structures, Computer Language Array Indexing, Generalized Hyper cubes, Array Division, Catenation, Lamination, Ragged Arrays.

Corrigendum (added in proof)

Negative array axis length results are inserted in this thesis as side comments and not as direct contributors to the main arguments. This is distinguished from negative arrays, which are main line items. The negative axis length results in the NERA case (equations 6.1.2 - 6.1.9, pp. 91-93) may be partially wrong or incomplete. Therefore for now, <u>all</u> negative axis length results should be considered preliminary and that entire topic as a whole should be considered part of Chapter 8, Implications for Future Research.

Preface and Acknowledgement

I have been developing the ideas in this thesis over many years in fits and starts. I wrote up a series of very preliminary chapters covering some of the decomposition material during the late 1980s and early 1990s as a personal time project while at IBM (I lucked out and got one of the first IBM PCs in a personal purchase lottery). I worked in the APL development group from 1970 to 1974 in Philadelphia, and in the APL R & D group at IBM Research from 1985 to 1987. This thesis work was done mostly after I left IBM in 1993. See the reference below to the 1987 IBM-APL-ITL (Interdivisional Technical Liaison) meeting.

To the best of my knowledge, these ideas are uniquely mine developed on my own time with my own resources. Although I received personal encouragement from Adin Falkoff, my manager at the time, my subsequent management was not so encouraging, as my required research was always set in other directions and there was no interest in theoretical data structures. I never got to work explicitly on this topic during the work days except to prepare an IBM ITL presentation while reporting to Adin Falkoff (see below). All of the work on this thesis was done long after I retired from IBM.

Many people have influenced me in general. I can't remember them all nor can I isolate out which attitudes of mine came from which person or from interactions of which group of people.

What does stand out in my mind are the lunches of the APL Development Group. These were informal meetings, held at lunch time around a large conference table in the IBM Philadelphia Scientific Center during 1970-1974. While we ate our lunch, we bandied around various ideas. Most topics were related to extensions of the APL language. Some were about applications of APL. Some were about more fundamental mathematics or array ideas. Some were about the politics and the sociology of APL or IBM product development. Others were about general politics. All were stimulating.

I also was influenced by the larger IBM APL community through the various IBM ITL (Interdivisional Technical Liaison) meetings. I did present a very preliminary version of the first level decomposition (of regular cubes) at one of these meetings in 1987.

I have also been influenced by the larger APL community via the International Conferences, some of which I did attend, and at which I did present papers on other topics.

When I was at the IBM Cambridge Scientific Center, I helped an MIT student (Jeremy Bloom) with his thesis that involved a substantial APL implementation. This brought me into contact with the MIT National Bureau of Economic Research. Their TROLL language was, as I remember it, a FORTRAN based language for econometric modeling that incorporated many callable routines that I viewed as essentially the APL array primitives under other names. This enlarged my views of array processing.

To name names is dangerous not only because I will forget some, but because many conversations in the APL or IBM Research communities were intellectually adversarial so that the person might not want to be associated with this work.

Nevertheless, I remember interacting with Yon Bard, Phil Benkhard, Jim Brown, Graham Driscoll, John Gerth, Adin Falcoff, Dave Grossman, Ken Iverson, Dick Lathwell, Dan Ling, Gene Mcdonald, John McGrew, Trenchard More Jr., Don Orth, Ray Polivka, Dave Rabenhorst, Cory Scutt, Howard Smith, Beth Tibbits, and many others whose names escape me now.

None of these interactions were directly about this work except for an informal presentation to Dan Ling about my personal work. He was not encouraging, but I did find the presentation useful in orienting me to better explain the uses of null arrays.

I had many conversations with Trenchard More Jr. while we were at the IBM Cambridge Scientific Center. We talked mostly about his emerging "Array Theory". I think I helped in part by introducing the idea of pointer or address name as an array element [24]. The important point concerning this thesis is that none of our conversations were about these ideas or about purely structural approaches to arrays. As I see it now, this work is orthogonal to Array Theory, intersecting only at the "origin", i.e. our use of the scalar as null array of 0 dimension.

More recently, at Pace University, my thesis advisor Fred Grossman annoyingly kept wanting more than a coherent, novel, well worked out theoretical framework. He kept asking "so what?", meaning that new knowledge, in and of itself, is not enough. We disagree. However, to placate this old friend and keep him quiet, I found some applications for him. The applications discussed here are merely indications of major independent investigations that could be undertaken. I don't consider the four application sections complete expositions, but merely indications of how my ideas can be profitably applied in other areas.

Fred and I come out of the NYU Courant Institute Of Mathematical Sciences - Göttingen applied mathematics tradition of David Hilbert and Richard Courant. I remember Lipman Bers pointing out that there are an infinity of true theorems but one should only work on those that have consequences. My theorems have many consequences.

I also remember reading the number theorist G. H. Hardy's proud comment in his Mathematicians Apology [15], that he never worked on any mathematics that had an application. So, this disagreement goes way back. Eugene Wigner [28] pointed out the "unreasonable effectiveness of mathematics" in physics. I think Truth will eventually find a use. Even Hardy has a central result in genetics (The Hardy-Weinberg Law) named after him, because he derived it, theoretically, in passing.

I must thank Michelle Lang, Anne Campbell and the staff of the Pace libraries for excellent support in retrieving most of the references I used here.

Special note should be made of the opportunities afforded to me by Dean Susan Merritt, Dean of the Pace University School of Computer Science and Information Systems. She was instrumental in my undertaking this entire program by making me an "offer I couldn't refuse" - it was so good.

In previous studies, years ago, that prepared me for this work, I received substantial help from Claire and Harold Leegant, my inlaws. During those studies I received encouragement from Prof. Peter Lax of the NYU Courant Institute.

Over the years, early on, I have to remember my parents and my Uncle Harry for encouragement and guidance. I also want to point out the many and various forms of help I got from my wife Dr. Judith M. Frank, including encouragement and editing. She didn't complain about the infinity of chores not done.

Table of Contents

List of Figures

List of Tables