

専用計算機

牧野淳一郎
天文シミュレーションプロジェクト
理論研究部
国立天文台



理論懇シンポジウム 2010/12/21

今日の話の構成

- 林先生のスタイル
- 自己重力系の物理
- GRAPE
- (林グループの研究活動の定量分析)

林先生のスタイル

昨日の杉本さんの話から (牧野のメモからなので適当)

extend problem as wide as possible
then concentrate to the central problem
avoid unclear assumptions
return to physics
construct a system from elementary processes
construct a whole story

自己重力系の物理

(牧野にとっての) 研究の原点

Hachisu and Sugimoto 1978, PTP 60, 123

- 等温ガス球
- 仮想的な球形断熱壁
- エントロピーの変分だけで議論

(「林先生のスタイル」とはなんか違う気がする)

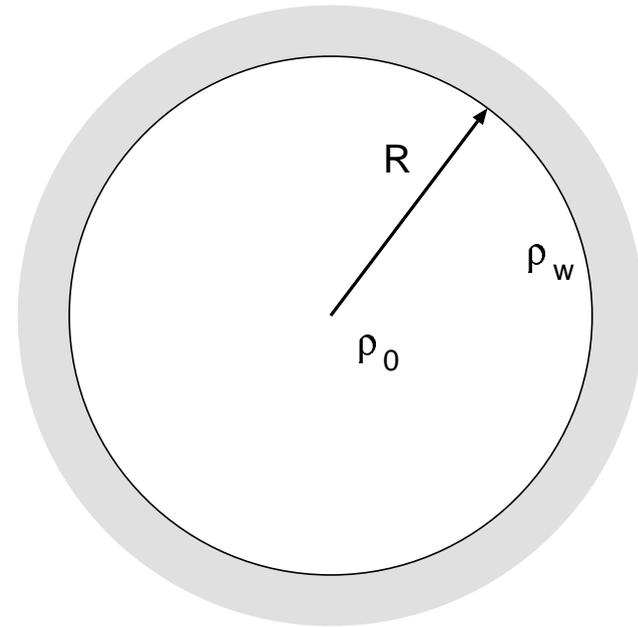
HS78

問題設定:

半径 R の中に質量 M のガスがあり、中心密度 ρ_0 と壁のところでの密度 ρ_w の比が D であるとする。

$D = 1$: 重力なし (温度無限大)

$D = \infty$: 特異解 $\rho \propto r^{-2}$



密度比とエネルギー

$D \rightarrow \infty$ の極限で、全エネルギーは $-1/4$ に振動的に収束

全エネルギーの最初の極小: $D = 709$
HS78 の解析: 熱の再配分に対するエントロピーの2次の変分を求める

- 1次の変分は0(等温だから)
- 2次の変分:普通は負に決まってる(等温状態=エントロピー極大)
- $D = 709$ では2次の変分が0になる摂動がある。 $D > 709$ では正もある

「重力熱力学的不安定」

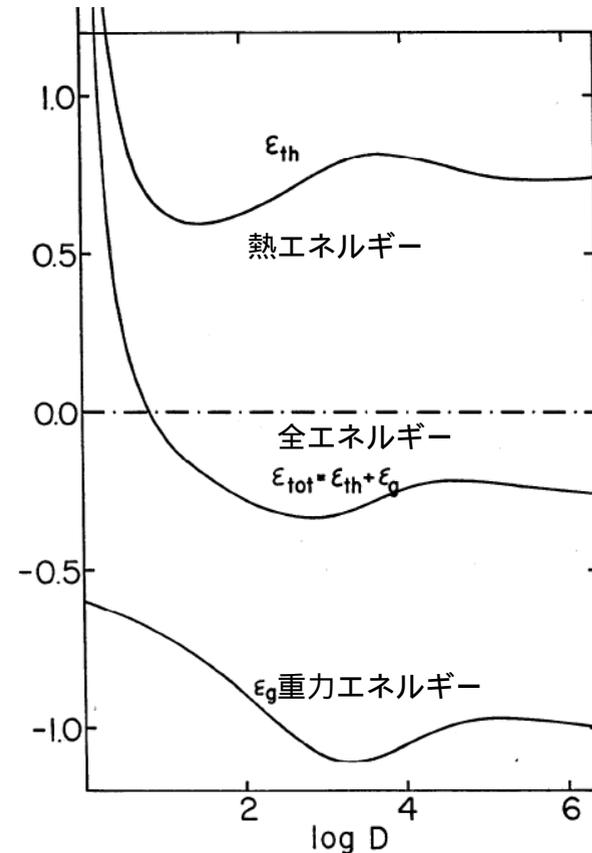


Fig. 2. Isothermal gas spheres in gravitational equilibria. The non-dimensional thermal $\epsilon_{th} = (3/2)\theta_0$, gravitational ϵ_g and the total ϵ_{tot} energies are shown against the central condensation D .

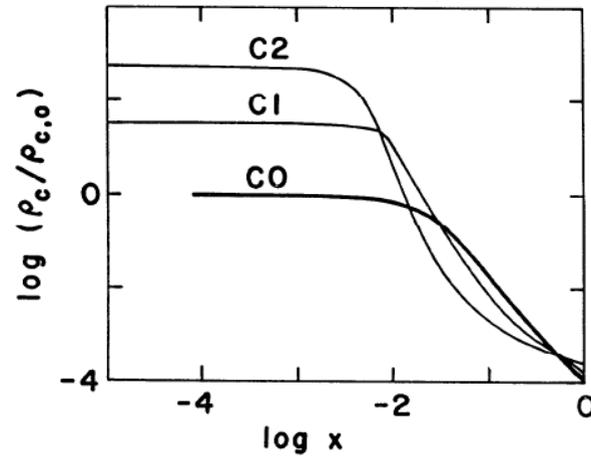
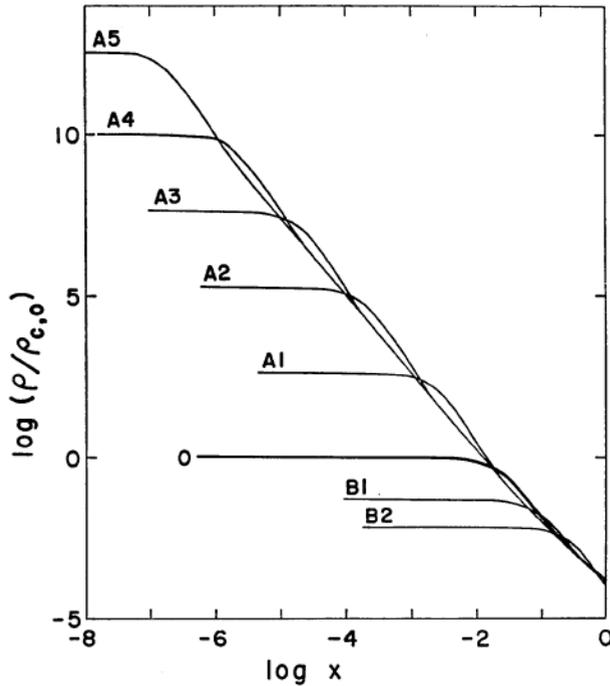
ちなみに

HS78: 本当に線形化した応答の固有値問題を解いた

$D = 709$ で中立安定であることは、固有値問題を解かなくてもわかる:
全エネルギーが極小 = エネルギーを変えないで D を微小量変化させる
ような摂動がある。(Antnov 1961, Lynden-Bell and Wood 1968)

非線型領域での進化: Hachisu, Nakada, Nomoto and Sugimoto 1978

HNNS78



熱伝導率 $K \propto T^{2.5}$

熱伝導率 $K \propto T^{-1}$

Lynden-Bell and Eggleton (1980): Self-similar solution,
熱伝導率 $K \propto \rho T^{-1/2}$ (恒星系の緩和過程を模擬)

→ 左のケースと定性的に同じ解 (べきはちょっと違う)

HNNS78 の教訓 と牧野が思うもの

- 構造形成の理解には数値計算が必須
- microphysics の仮定が激しくアレでも正しい答になることがある。
それには理由がちゃんとある
 - $T^{2.5}$ と $\rho T^{-1/2}$ はだいぶ違うような気がする
 (“This is not the physical case.”)
 - 解の振る舞いはほぼ同じ
 - self-similar 解にそって考えると、どちらも中心でタイムスケール短い。

HNNS78 の教訓 と牧野が思うもの

- 構造形成の理解には数値計算が必須
- microphysics の仮定が激しくアレでも正しい答になることがある。
それには理由がちゃんとある
 - $T^{2.5}$ と $\rho T^{-1/2}$ はだいぶ違うような気がする
 (“This is not the physical case.”)
 - 解の振る舞いはほぼ同じ
 - self-similar 解にそって考えると、どちらも中心でタイムスケール短い。
- 偉い先生も色々間違える。

とはいえ

林先生のスタイル:

- あらゆる物理に通暁する
- 全部いれて計算する
- 結果の本質を理解する

とはいえ

林先生のスタイル:

- あらゆる物理に通暁する
- 全部いれて計算する
- 結果の本質を理解する

杉本先生のスタイル:

- あらゆる物理に通暁する
- 本質だけをいれて計算する
- 結果の本質を理解する

とはいえ

林先生のスタイル:

- あらゆる物理に通暁する
- 全部いれて計算する
- 結果の本質を理解する

杉本先生のスタイル:

- あらゆる物理に通暁する
- 本質だけをいれて計算する
- 結果の本質を理解する

どちらも凡人には無理

とはいえ

林先生のスタイル:

- あらゆる物理に通暁する
- 全部いれて計算する
- 結果の本質を理解する

杉本先生のスタイル:

- あらゆる物理に通暁する
- 本質だけをいれて計算する
- 結果の本質を理解する

どちらも凡人には無理

一体どうすると？

というわけで

もうちょっと違うことを考えないといけない

- 人が(まだ)やってないことをやる
- 人にできないことをやる

という話も昨日あったような気がする

非線型な構造形成の良い(悪い?)ところ

- 大抵のことについて、「こうなる」というまことしやかな「理論」がある
- 特に、大抵の場合に「今できているより大規模な計算はしなくてもよい」という「理論」がある
- そうするのは、結構全然間違っている

というわけで、

- 人が(まだ)やってないことをやる
- 人にできないことをやる

の一つ: 計算機にがんばってもらう

Fujii et al. 2010

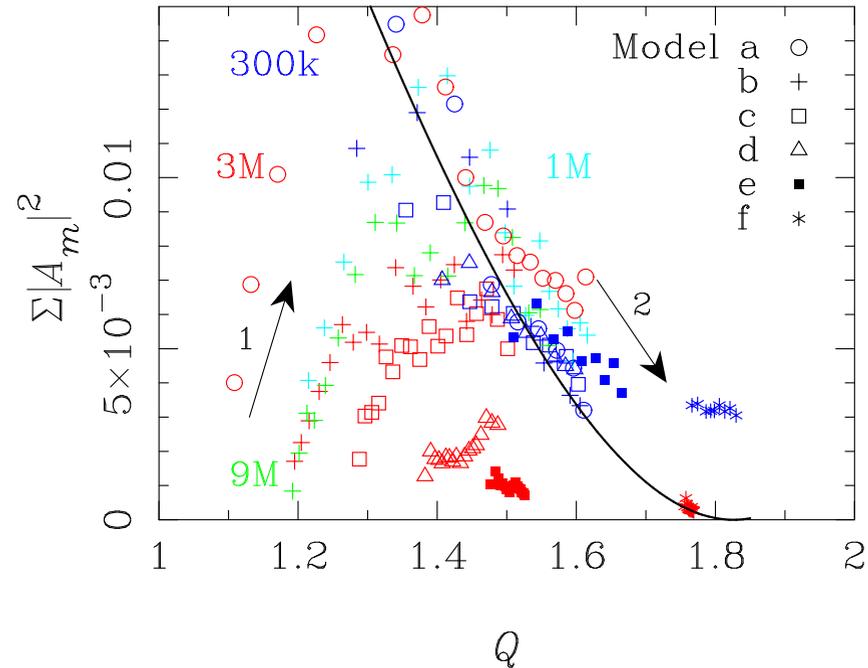
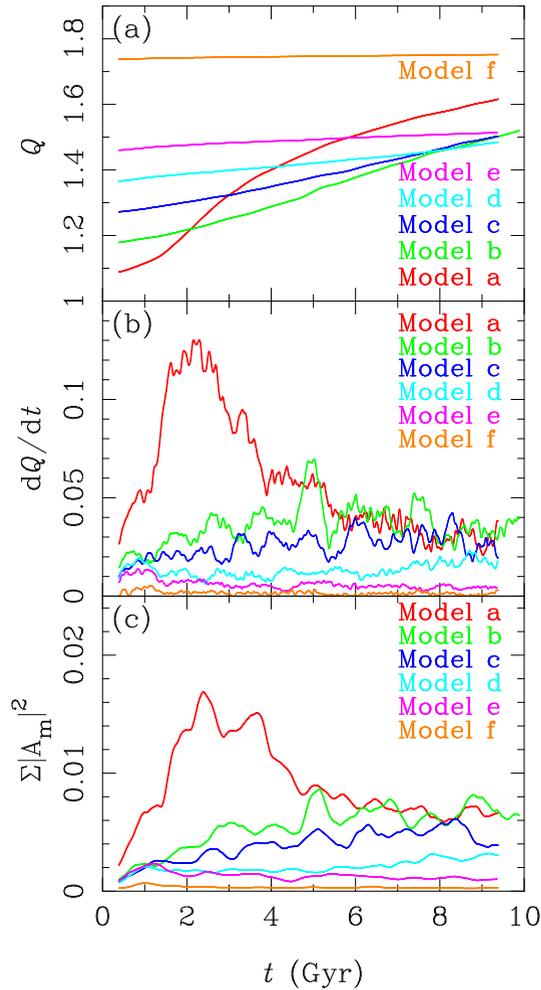
純粹に恒星だけの円盤のスパイラル構造

和田さんから藤井さんに「やってみない？」という話が

牧野は「そんなのすぐに消えるに決まってる」と思ってあまり良い顔を
しなかった。

結果: 10Gyrs たってもスパイラル消えない、、、

Fujii et al. 2010(2)



初期の Q が極度に小さくなければ、スパイラル構造は 100 億年以上持つ
 Q と腕の強さの平面上で収束
 「藤井トラック」

昔杉本先生に言われたような、、

(間違っていたらすみません、、、)

- 昔は、研究は、1/3は何をするか考え、1/3は計算して、1/3は論文書くものだった
- 今は90%プログラム書いて残りで論文書いている

何をするか考えるところがなくなってる、、

- それではいかんので面倒なプログラムを作ったりしないようにする
- 簡単なプログラムでもどうせ90%の時間使うんだから、難しいことでも同じ

どっちもありだけど、「人のしないほうを」

人のしない程度

- 普通の計算機で動く簡単なプログラムを書く
- 普通の計算機で動く難しいプログラムを書く
- 使うの面倒だけど速い計算機で動く簡単なプログラムを書く
- 使うの面倒だけど速い計算機で動く難しいプログラムを書く
- 計算機作って簡単なプログラムを動かす
- 計算機作って難しいプログラムを動かす

計算機を作るところでほぼ他にだれもいなくなる。

計算機作ってまですることあんの？

これはまあ、プログラムでもなんでも同じ。

- 自分はどれくらい長い間同じことをするか？
 - 林先生: 10年くらい？ 他の人は？
- プログラムなり計算機なりは他のことにも使えるか？

近田提案

1988年、天文・天体物理夏の学校

88/11(2)

A. S.

1276

「計算機を使う，計算機の技術を使う」

近田義広 (国立天文台・野辺山)

観山さんは理論の立場から見ると，現在の日本ではスーパーコンがどんどん使えるからイニシアチブをとる良い機会だと話されました。私は，その素晴らしい計算機を作り出す技術をうまく使いこなし，天文に応用すれば，計算機の利用者に止まる場合よりもっと大きな成果を期待できるぞ！ということをお話しします。

- 計算機買うんでなくて作れ!

(昨日, 今日, 明日)

先ず, M1, M2の若い人達に話す時の定石の昔話から始めます. 約20年前の1960年代の末, 三鷹の天文台に口径6mのミリ波望遠鏡が作られ, 後の野辺山の望遠鏡の雛形が生まれた時, その制御とデータ収集はOKITAC4300と言うミニコンが受け持ちました. そして今, 6m鏡は水沢に運ばれてVLBIに使うために化粧直し中. 今度の制御用計算機は日電のパソコンPC9801VMです. 表1に見るように, この20年間に記憶容量も演算速度も約4桁良くなっています. つまり, 同じ予算でも10年たてば100倍の仕事をさせられるようになったと言うことです. これは天文台と言うローカルな世界だけのことではありません. 事実, 図1〔1〕の並列計算機の並列度の推移にみるように10年で100倍というのはかなりよいestimateです. 但し, これは重要なことなのですが, 同一の構造, 同一の使われ方, つまり同一の製品系列の中での進展はこんなに早くありません. 例えば野辺山の汎用計算機についてこの4年間の価格/性能比を見るとせいぜい2倍しか良くなっていません. つまり計算機(の技術)の使い方を常に見直して, 自分の問題に適した構造の機械を使う(作る)ようにし, かつその上で解き易いように問題を立て直すことがないと「10年で100倍」を生かした人に比べ10倍分は立ち遅れることになりま

- 計算機は10年で100倍速くなる
- 同じ製品系列だと10倍しか速くならない

全体の値段 \propto 性能 \propto 装置の大きさ

でもあります。表2に見るように上記3者の比例関係は大体成り立っているように見えます。しかし本当に価格/性能比は一定なのだろうか？ もしそうなら利用者は利用者に徹し、計算機システムの構成に気を使ったり、自分で情報処理機を作ったりしようなどとせず、与えられたお金でCPU時間を買っていれば良いことになります。誰もが平等にお金の額に相当する「性能」を手にするのだから、残された仕事は金を取ってくることだけです。

そうではありません。汎用の大きな計算機は、出来るだけ広い範囲の利用者の平均的な要求を同時に満足させる「性能」を考えて作ってあります。ですから汎用「性能」を考えた場合は上の比例式は成り立っていますが、特定の範疇の性能を取り出すと比例関係は成り立ちません。だから、問題は自分の課題に対する「性能」とは何か？それを一番安く（容易に）手に入れるにはどうしたらいいか？です。例えば野辺山クラスの計算機1台とEWS(engineering work station)400台とどちらがいいか？もし演算速度だけを問題にするなら、EWSを400台の方が良いでしょう。またFFTの速さだけなら野辺山のFXはスーパーコンピュータに比べて2桁以上速くて（表3）値段は1桁安いので、その方がいいに決まっています。汎用か専用かでこんなに性能/価格にひらきがでるのは一体どうしてなのでしょう。

- 専用化すれば2 - 3桁価格性能比をあげられる

具体的イメージをつかんでもらうために、例として図2にパソコンにつないでN体問題の重力を計算する回路の概念を示します。10MHzクロックで間口がパラレルのバイプライン式で32bit固定小数点で計算させるとすると、市販のICで200個、0.1立方メートルの高で、250MOPS位のスーパーコンピュータ並みの性能が得られるでしょう。大体、1万問題を1万ステップ1日に計算できる勘定です。値段としては自作で400万円位でしょう。もし億円オーダーの金をかけて、大型計算機につなぐようなバンとしたものを作る場合は、高性能をねらって専用ICを作り、IC5000個、4byte浮動小数点演算、20MHzクロック、20立方メートル、100GFLOPS、10万問題を4万ステップ1日に計算できる。(但し、近田電子製作所の見積もりは甘いと言う声もあることを付け加えます)。

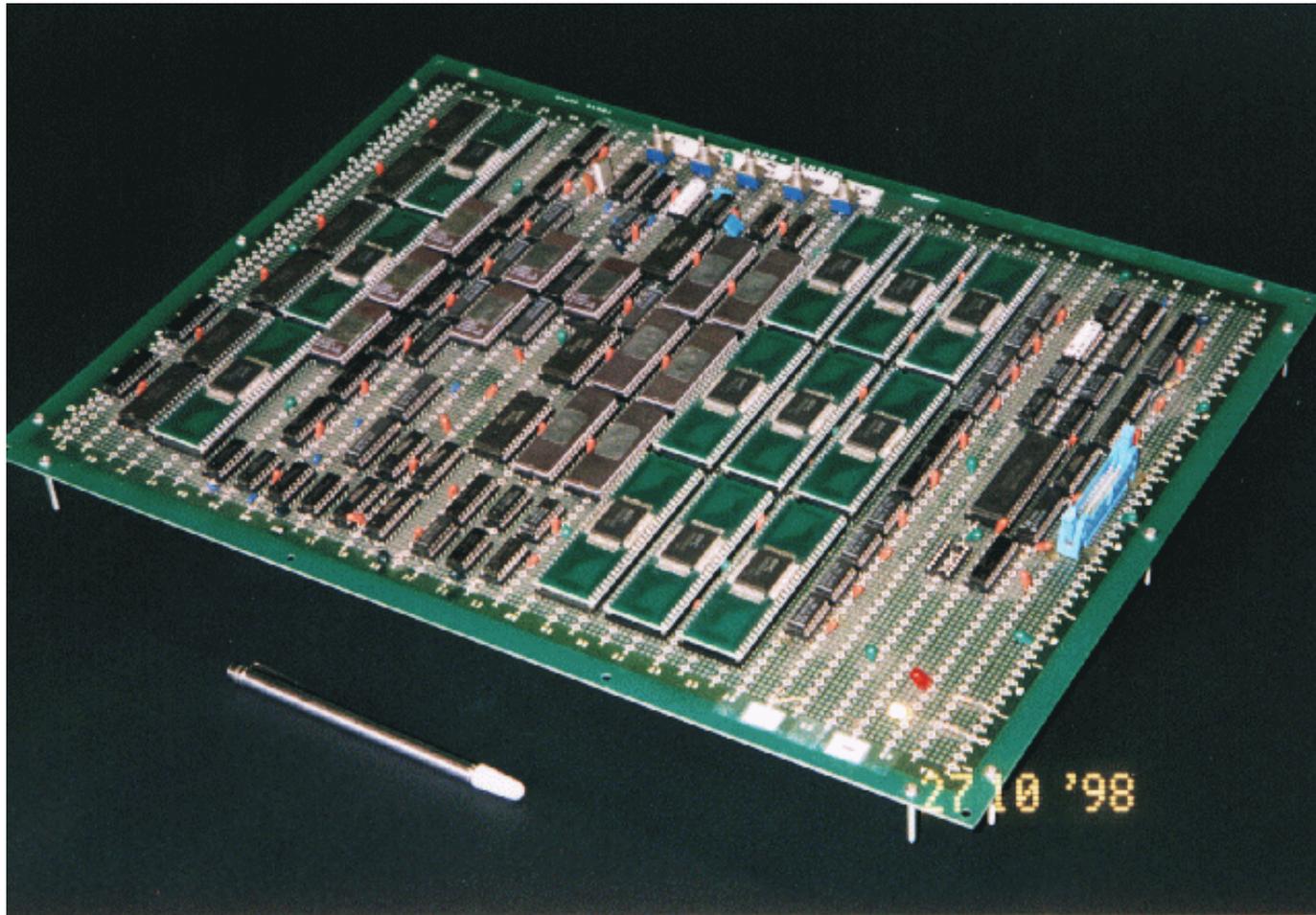
- 多体問題なら 400万円 250 Mflops
- 数億円なら 100 Gflops
- 「但し、近田電子製作所の見積もりは甘いという声もあることを付け加えます」

杉本さんの反応

- 11月(だったと思います)に近田さんの話を聞きに野辺山に
- 川合先生(慶応)、牧野、伊藤君が同行
- 近田さん、奥村さん他で話を。

「近田さんが色々教えてくれそうだから、やってみよう」

GRAPE-1 (1989)



GRAPE-1 開発プロセス

- 88年12月くらいまでは、伊藤がデジタル回路の勉強、平方根回路の検討とか
- 89年2月前後に、回路規模検討、ROM テーブル、対数フォーマット等検討
- 6月くらいまでに概略仕様、使うチップ等決めたような気がする
- 8月終わりにはハードウェアができていた

開発はどんなふうだったか



詳しくは伊藤君の本をどうぞ。

ハードウェアは私が知らないうちにできていた (伊藤君が作った) ので良く知らない。

GRAPE-1 と近田原提案の違い

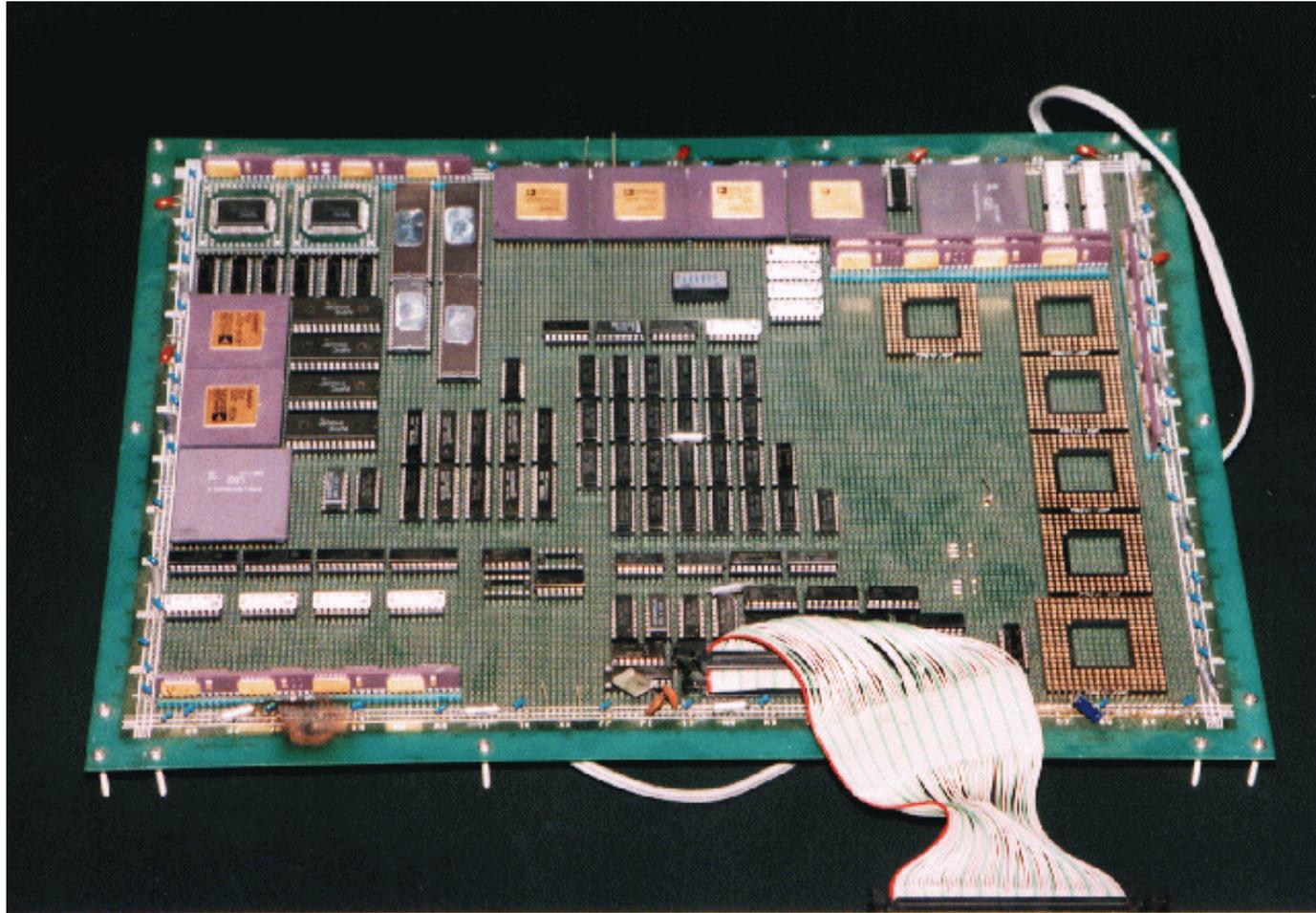
- パイプライン構成は基本的に同じ
- 大きな違い: 数値表現の形式
 - 近田さんの想定 (多分): 固定小数点 32 ビットで基本的にはそのままいく。2 乗すると 64 ビットになって、 -1.5 乗の結果は 96 ビットになるけど、、、
 - 当時の素直な解決方法: 85 年くらいからでてきていた、浮動小数点計算専用 LSI を使う。
 - GRAPE-1: 途中を、8 ビット対数表現にする。

ハードウェアコストを近田電子製作所見積もりより 1 桁下げた!

何故対数表現？

- 8 ビットに収めなかった: 512kbits の ROM で 2 項演算が実現できる。 $(x^2 + y^2)$ とかが石 1 つですむ)
- 最初と最後だけ固定小数点形式にすれば、途中は全部対数形式でも問題によっては十分な精度がでる。
- 牧野は学部 3 年生の時に学科の特別講義で石黒さん (多分) が野辺山 FX の話をして、数ビットで計算しても大丈夫というようなことをいったのをおぼえていたような気がする。
- 無衝突系の場合、2 粒子間の力の必要な精度は低い。多数の粒子があることで統計的に高い精度が出れば十分 (ということを経験的に証明できた)。

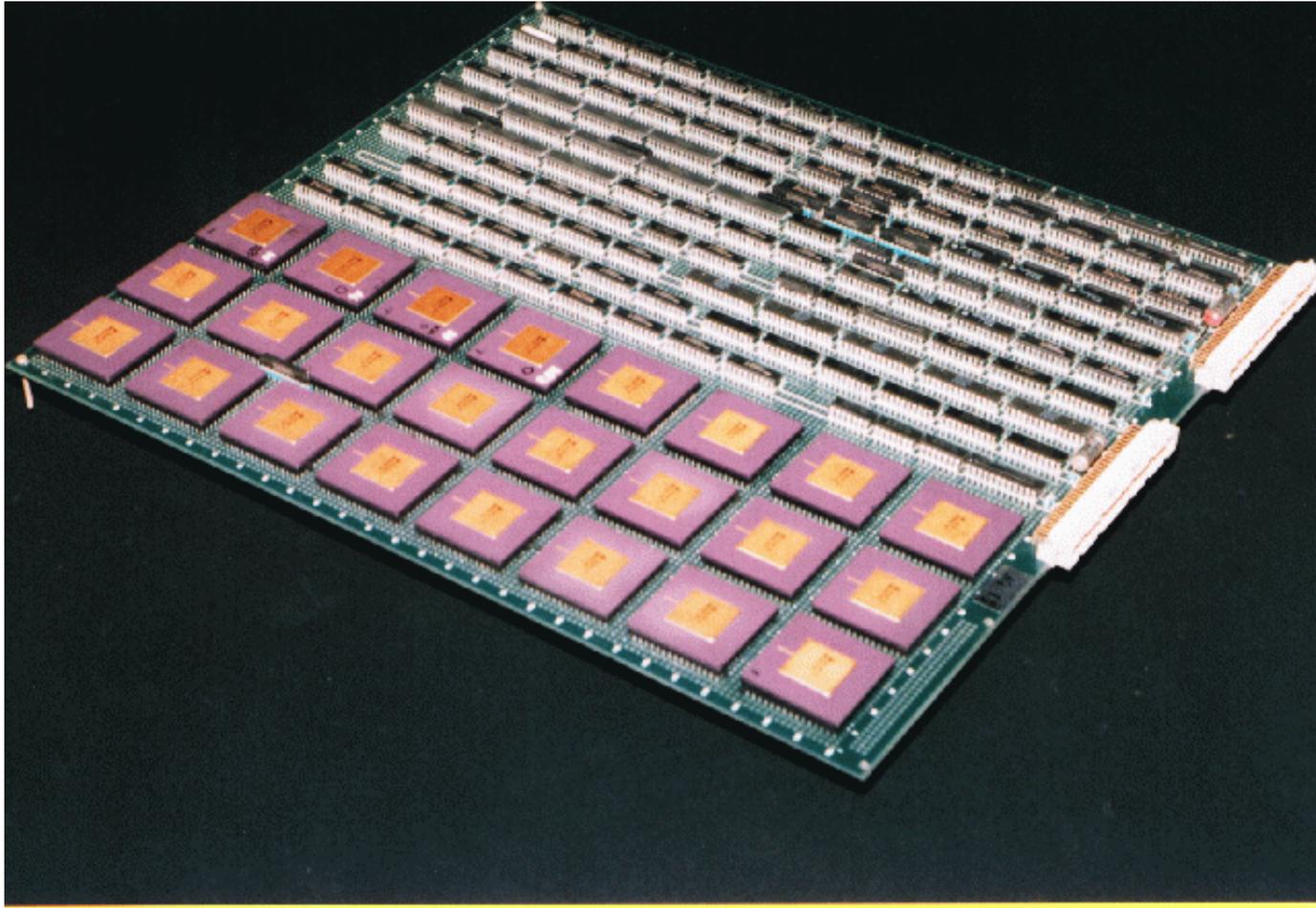
GRAPE-2(1990)



GRAPE-2 の中身

- 8ビット演算とかは止めて普通に浮動小数点演算 (倍精度は最初と最後だけ)
- 高精度が必要な問題も色々あるため
- 40Mflops
- 戎崎、伊藤、牧野
- 詳しくは伊藤君の本を参照

GRAPE-3(1991)



GRAPE-3

- 大雑把には GRAPE-1 パイプライン (但し、大きな ROM は使えないので対数での加算の実装は全く違う)
- 仕様決定、シミュレータ (C で記述) は牧野
- 論理設計以降は富士ゼロックス (橋本、富田)
- SCS Genesil で設計
- ファブは NS. $1\mu\text{m}$
- ボードは奥村
- 1千万ちょっとの予算で 20Gflops を実現。

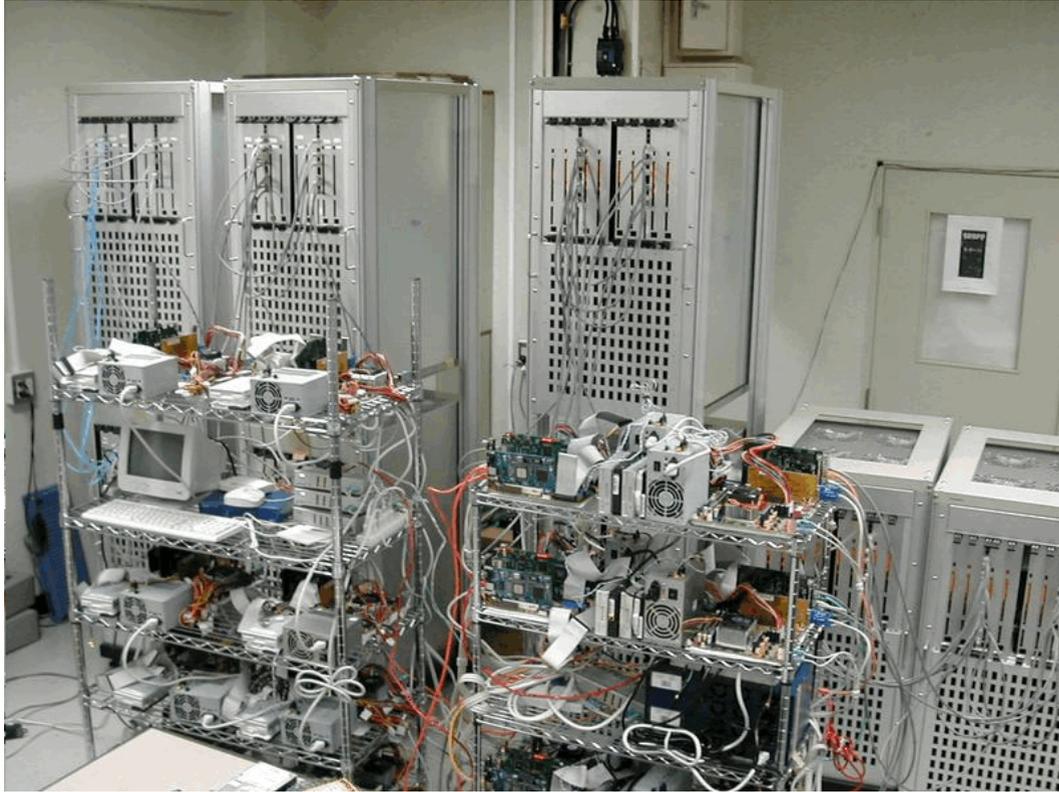
GRAPE-4(1995)



GRAPE-4 概要

- 48 個プロセッサチップがのったボード 36 枚
- 1 チップ 20 演算、32 MHz 動作で 640 Mflops
- LSI logic 1 μ m スタANDARDセル
- プロセッサチップは泰地、「予測子」チップ牧野

GRAPE-6 system



2002年の 64 Tflops システム

4 ブロック

16 ホスト

64 プロセッサボード

他の GRAPE 型機械

- 1991 GRAPE-1A : GRAPE-1 の細かい改良。設計:福重
- 1992 GRAPE-2A : 最初の MD 用 GRAPE
設計:伊藤・福重
- 1992 HARP-1: 小久保君設計。 Hermite 公式用
- 1993 GRAPE-3A: 商業版
- 1996 MD-GRAPE: 最初の MD 用 GRAPE チップ
チップ設計:泰地
- 2001 MDM (MDG2): 理研・戎崎グループによる。 75T
- 2006 PE (MDG3): 理研・泰地グループによる。 1P
- 1992- MD-Engine: 富士ゼロックス、大正製薬 (現在 NEC)

GRAPE によるサイエンス

他の人にできないことができる

- 星団進化
- 惑星形成

微妙

- 宇宙論
- 銀河形成

他の GRAPE 関係専用計算機

- 1991 DREAM: 大規模偏微分方程式計算「専用」計算機。ハードディスクを主記憶に、という発想。ディスク 1 台の試作程度まで。大野、牧野、戎崎。
- 1993 ZEBRA : Radiocity 法専用計算機。成見、戎崎
- 1995 General: LU 分解専用計算機。市販チップで構成。清木、戎崎、泰地、牧野。
- 1994-? SPH 専用計算機。横野他
- 2002? MACE: LU 分解専用計算機。泰地、戎崎

GRAPE 以外はどうだったか

素晴らしい成功だった、とはいいがたい、、、
そうなった理由(多分):

- 使ってやるサイエンスから考えた設計までできたか？
- 「見積もりが甘」かったのでは？
- ...

GRAPE-DR

- 開発費の問題
- 解決の方向
- 現状

開発費の問題

- 半導体の進歩に比例して開発費用も増大

年	計算機	チップ初期コスト	設計ルール
1992	GRAPE-4	200K\$	1 μ m
1997	GRAPE-6	1M\$	250nm
2004	GRAPE-DR	4M\$	90nm
2010?	GDR2?	> 10M\$	45nm?

全体予算は初期コストの4倍程度必要。
15億の予算は天文専用ではとれない。

GRAPE-DR

- プログラム可能アーキテクチャ — GRAPE より広い応用範囲
- SIMD 並列計算機を1チップに。
- チップ性能: 単精度 512, 倍精度 256Gflops (500MHz で、、、)

基本的な計算モデル

$$R_i = \sum_j f(x_i, y_j)$$

を並列に評価。

- 2重ループ、一方について積算。
- y_j がなければ単純な並列計算。
- 行列乗算もできるように作る: Top500 狙うため

GRAPE-DR クラスタシステム



GRAPE-DR クラスタシステム

- 128-ノード, 128-カード システム (105TF 理論ピーク @ 400MHz)
- Linpack 実測性能: 39 Tflops@400MHz (100 ノード、理論ピークの半分)
- ホスト計算機: Intel Core i7+X58, 24GB メモリ (100 ノードくらい。残りは 18GB)
- ネットワーク: x4 DDR Infiniband
- 300 ノードくらいまでは冬なら今の部屋で。それ以上はもうちょっと広い部屋 (電源能力) 必要。

Little Green500 (2010/6)

Listed below are the Little Green500's Top 10 most energy-efficient supercomputers in the world as of June 2010.

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	815.43	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	28.67
2	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
2	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
2	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
5	536.24	Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw	BladeCenter QS22 Cluster, PowerXCell 8i 4.0 Ghz, Infiniband	34.63
6	530.33	Repsol YPF	BladeCenter QS22 Cluster, PowerXCell 8i 3.2 Ghz, Infiniband	26.38

- ベンチマークは Top500 と同じ
- 消費電力当りの性能でランキング
- ちょっと小さいめのシステムまであり、のリストで1位
- 何故かNHK まで取材にきた。

Green 500, Nov 2010

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
<u>1</u>	1684.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype	38.80
<u>2</u>	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows	1243.80
<u>3</u>	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core, NVIDIA C2050, Infiniband	36.00
<u>4</u>	828.67	RIKEN Advanced Institute for Computational Science	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	57.96
<u>5</u>	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
<u>5</u>	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
<u>5</u>	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
<u>8</u>	740.78	Universitaet Frankfurt	Supermicro Cluster, QC Opteron 2.1 GHz, ATI Radeon GPU, Infiniband	385.00
<u>9</u>	677.12	Georgia Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5660 2.8Ghz, nVidia Fermi, Infiniband QDR	94.40
<u>10</u>	636.36	National Institute for Environmental Studies	GOSAT Research Computation Facility, nvidia	117.15

??? Where is GRAPE-DR ???

Green 500, Nov 2010

- The performance we submitted: **1474 Gflops/W**
- It **should be** #2 in the list
- Somehow we were not listed
- Little Green 500 list is not even released yet
- Green 500 people are “working” on this (according to them)

GRAPE-DR は上手くいったか？

- 言い訳: 牧野が天文台に移ってやたら忙しく、、、
- 技術的には、完全に専用化した GRAPE に比べて汎用計算機に対するメリット小さい
- 結構似たアーキテクチャの GPGPU がほぼ同時期に発生した
 - トランジスタ効率で10倍しかよくない

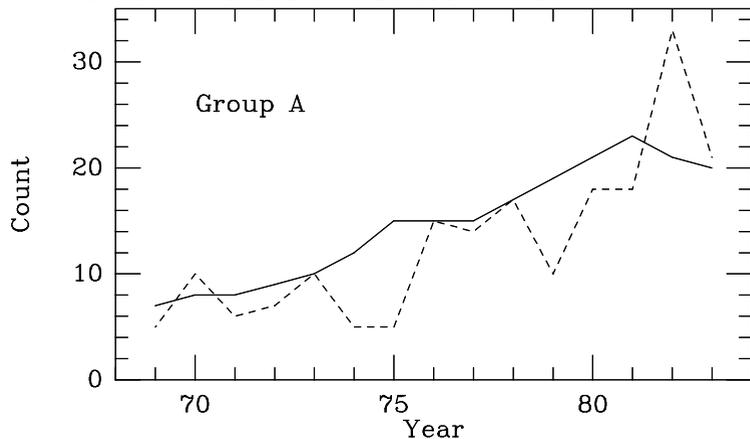
GRAPE のような大成功では今のところはない。まあ中成功くらい？

終わりに

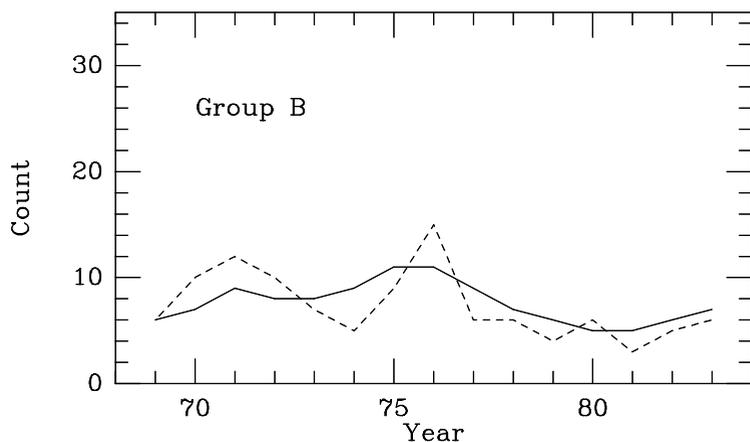
- 専用計算機は、要するに「人のやらないことを」という話
- 先生のやってないことを、というほうが正確かもしれない
- 若い人はもうちょっと違うことを考えてね (GPU を使うとかでなくて)

おまけ:林研の研究活動の定量分析

昔戎崎さんが言い出したネタ。戎崎さんは途中で理研にいったので、、、
 牧野、藤垣、今井、科学研究における研究グループの生産性と質的評価
 『年報 科学・技術・社会』 Vol. 6 (1997)



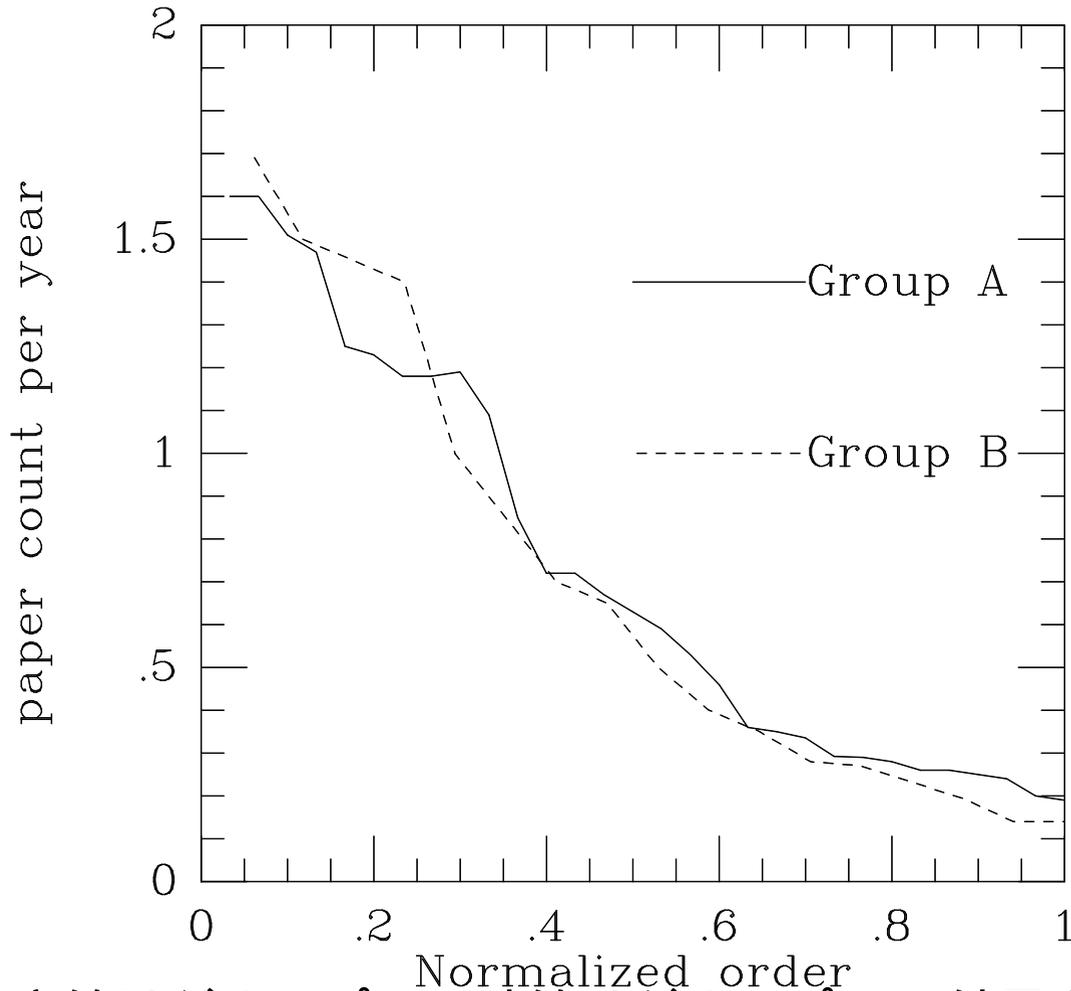
両グループの研究者数と各年の出版論文数の変化。実線は研究者数、破線は論文数を示す。



グループメンバーの生年

	グループ A	グループ B
1920 年代	1	2
1930	2	2
1940	11	9
1950	11	3

論文生産性の分布

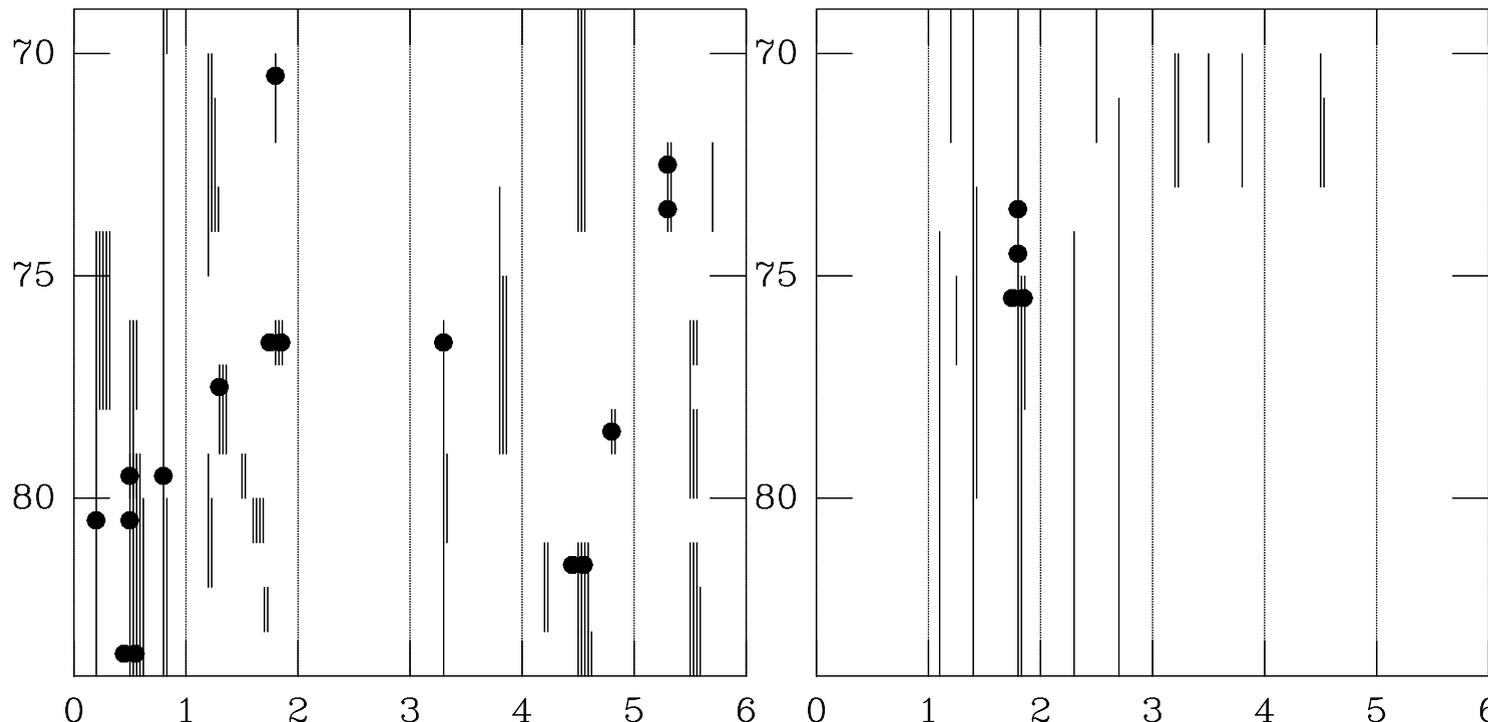


実線はグループ A、破線はグループ B の結果を示す。縦軸は研究者の 1 年当たりの出版論文数、横軸は研究者のグループ内での順位をグループ総人数で割ったものである。

被引用度の分布

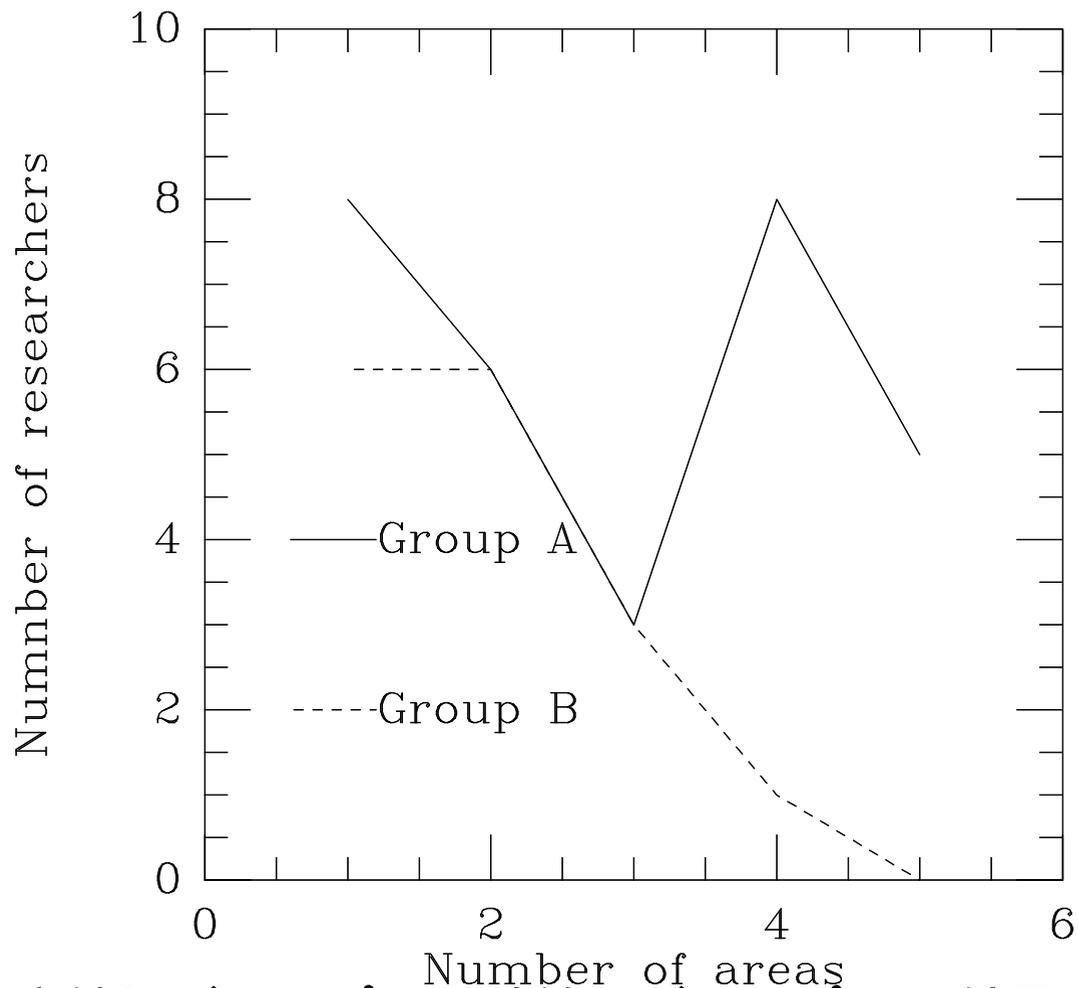
被引用回数	グループ A	グループ B
0—9	61.3%	52.7%
10—19	20.6%	34.5%
20—29	9.8%	9.1%
30—	8.2%	3.6%

論文を発表した分野の移り変わり



論文を発表した分野の移り変わり。左はグループA、右はグループBの結果を示す。横軸は表1の各分野を示し、縦軸は歴年である。線分は連続した内容の論文の存在を示し、本数は著者グループの人数を示す。なお、黒丸は被引用度30以上の論文を示す。

論文を発表した分野数の分布



実線はグループ A、破線はグループ B の結果を示す。

分野分類

我々の分類	A&A 分類番号
恒星・惑星系形成論	107, 131
恒星内部構造論	61-65, 120-129
太陽	70-80
銀河・銀河団	150-159
宇宙論	160-169
コンパクトスター	66, 67

A&A 分類

00X, 01X	色々	08X	地球
02X	数学	09X, 10X	太陽系
03X	装置	11X, 12X	星
04X	位置天文	13X	星間現象
05X	スペース天文学	14X	電波、X
06X	理論天文学	15X, 16X	銀河、銀河団
07X	太陽		