

基研の作るグローバルな研究環境

青木健一（基研）

基研の将来計画についての本日のシンポジウムでは、基研の将来像が議論されているが、もとより、多様な側面があることは言うまでもない。私は、その中で最も non-trivial な、つまり、個別大学や研究機関ではなく基研にこそ位置づけてほしい課題についてコメントしたい。

結論はタイトル通り、「基研はグローバルな研究環境を作ることに寄与すべきである」という事である。具体的なイメージとしては、誤解を恐れずに標語的にたとえて言うと、「Operating System “基研” Version 2 は UNIX で行こう」となる。ここでグローバルというのは、今やもちろん、全地球的なという意味である。グローバルな研究環境の中にも色々な要素があるが、話の後半では、その中でも最も単純明快な、それ故最も primitive で重要な研究情報の側面について述べる。

多様な側面がある中で、という言い方を最初にしたが、むしろ私は「グローバルな研究環境を作る」という事を広い意味で捉えれば、それで基研の果たすべき独自の役割は尽きていると考える。その中では、「固有の」とか「共同利用の」とかは対立する概念ではあり得ない。

1. 議論の出発点

研究者にとって大事（大切、幸せ）な事は何か、それは「研究において Active である事」に他ならない、という事を議論の出発点としたい。この Active であるという内容は、単に外から論文の数を数える事とは違う。この Active という言葉の持つ意味については、本シンポジウムでの高塚氏の話の中に適切なまとめがあったと思うが、私としてはさらに次の様な点を強調したい。つまり、これは研究者の側から見ての概念であり、といっても甘くするのではなくむしろ自己に厳しい自己評価を伴うものである。また、個々の研究成果に限らず、研究の Organization、つまり研究者の Society としての Activity を重視したい。そして、研究評価については、安易な評価の危険性は承知の上で、むしろ積極的に、成果に対する正当な評価を要求する権利、被評価権、を位置づけたい。安易な評価が横行すればするほど、研究者（Society）の自己評価能力が問われて来る。

2. データを少し

以前に、素粒子論分野に於ける世界の研究動向、Activity の分析を行なった¹⁻²⁾。その中から二つの例を引く。以下のデータは文献データベースを利用して計算したものであり、

Fig.1

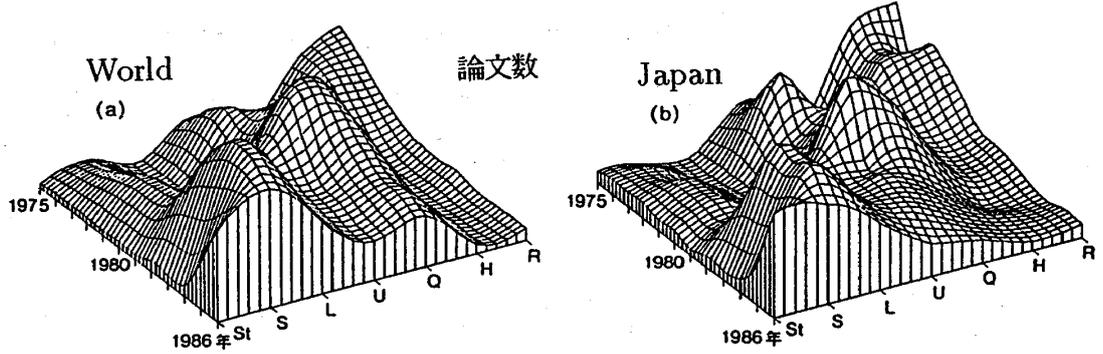


Fig.2 Shaded: Japan, Dotted: R

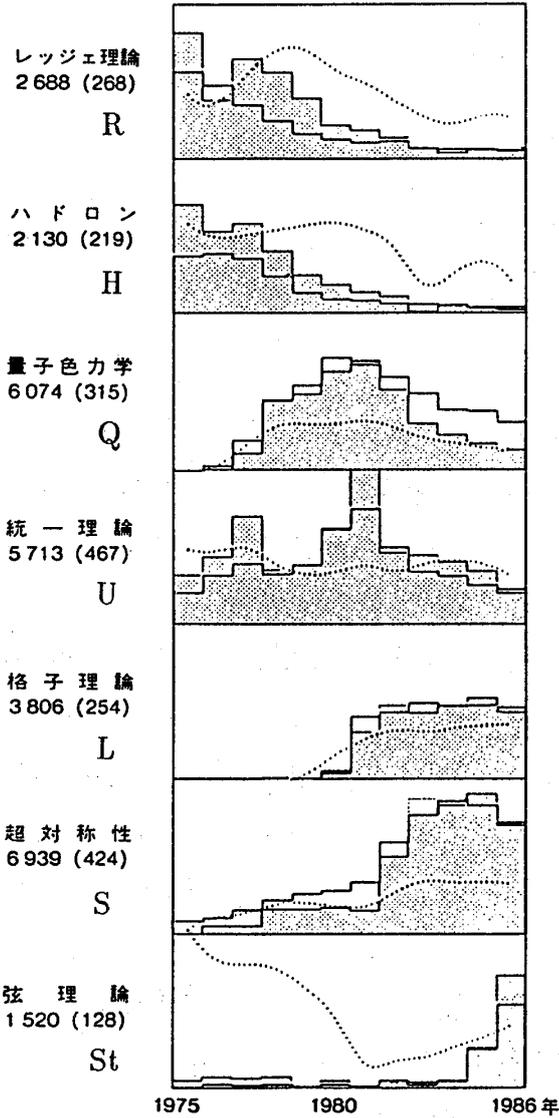


Fig.3

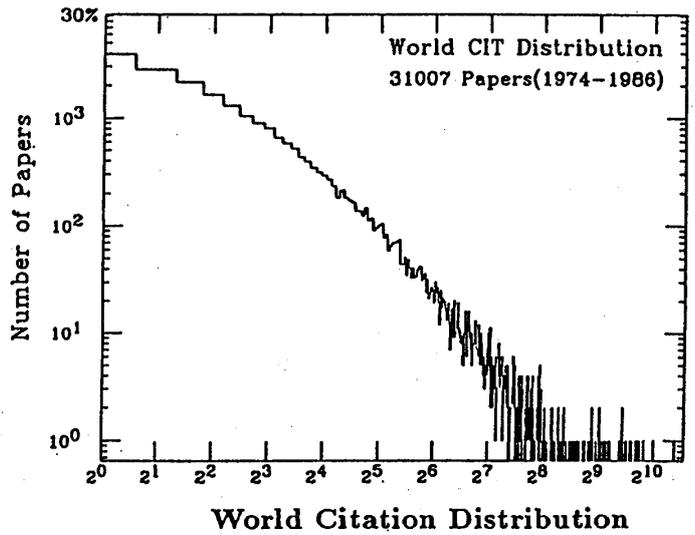
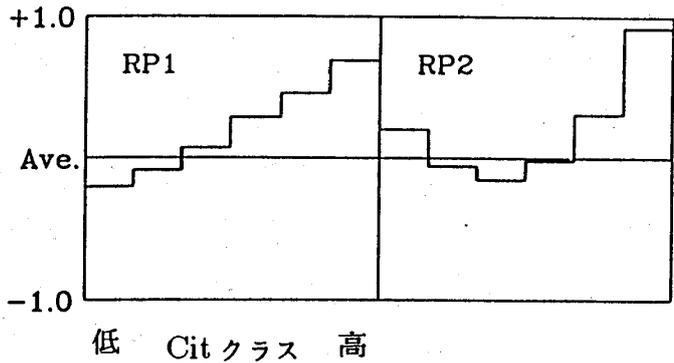


Fig.4 Eigenfunctions for RP1, RP2



バルクな性質 (平均値) を記述している。分散が問題であるが、一方、悉皆データなので、(サンプリングではない)、集団間の平均値の差には大きな意味がある。

a 流行

学問の進展にいわゆる「流行」があることはよく知られている。これには、安易な流行追随とその弊害という文脈で語られるべき事が多いのも事実であるが、むしろ、世界の研究者

の総合的で広い視野による合理的な研究方針設定の必然の一致として避けられないという面もある。実際、多人数による力の集中は、大きな壁を越えるのに必須である。するとむしろ課題は、originality の高い、流行を作るような仕事をどの様にして生み出すか/organizeするか、新しいアイデアの芽が育つ研究環境をどう作るか、ということにある。

データベースを利用して、1975年から1986年のストリング全盛のころまでの素粒子論のテーマの変遷を見ると、おおきなうねりが動く様子がよく分かる (Fig.1)。これの断面を世界全体 (W) と日本の部分集合 (J) で比べ、個々のテーマ毎に「世界でのそのテーマの割合に対する日本でのそのテーマの割合の比 ($R = \text{Rate}(J) / \text{Rate}(W)$)」の年変化を調べると (Fig.2の点線)、一部のテーマを除いて、Rは山型の変化を示す。この事は、日本の部分集合は、立ち上がり少し遅れるが、ピークでの「専念度」は高く、ピークを越えた後の逃げが素早い、事を意味する。「早逃げ」というのは結局、立ち上がり始めた次の流行へと目が移ることに過ぎないから、日本の部分集合はまだまだ世界を大きな熱溜として感じている事になる。

b 論文の Citation と寿命

もう一つは、Citationの分析である。論文の獲得した Citation 数の分布を見ると (Fig.3), 分布 $P(n)$ は $P(n) = n^{-1.6}$ でよく近似される。これは、 $2^{1.6} = 3$ なので、「2倍よく引かれる論文は3分の1しかない」というスケーリング則になる。Citation というのは集団現象の面があり、対数 (Log) が適切なスケールであることが分かる。

そこで、Citationの対数スケールで論文をクラス化し、各研究機関毎にクラス分布割合を計算し、その統計分布を見ると、最初の2つの主成分で分布がよく記述される (Fig.4)。第1成分 (RP1) は単純に、高いクラスの論文が多いかどうかを表現している。第2成分 (RP2) は、ホームランか三振かというタイプと打率は高いがシングルヒットが多いというタイプを区別する方向である。この2次元分布は三日月型の分布をしているが、驚くべきことに、世界中の大きな研究機関をそれぞれの地域 (A:アメリカ、W:西ヨーロッパ、E:ソ連と東ヨーロッパ、J:日本) に分けると、地域毎に大きく離れたクラスターを形成している (Fig.5)。この事は第一に、この分野でのアメリカのトータルな意味での優位性を示しているが、それだけではない事情をこの分析から引き出す事ができる。

ある論文がどのくらいの期間にわたって引かれ続けるか (論文の寿命) もまたその論文の価値の Citation とは違う別の指標である。この論文の寿命と Citation との相関を見る。そこで、各 Citation クラス毎に、論文が書かれてからの経過年数 (横軸) による各年での獲得 Citation 数の変化を計算する (Fig.6)。それぞれのクラス毎に特徴的な寿命で各年毎の Citation 数が減っていくことがよく見える。Fig.6aはA地域の例であるが、Citation

Fig.5

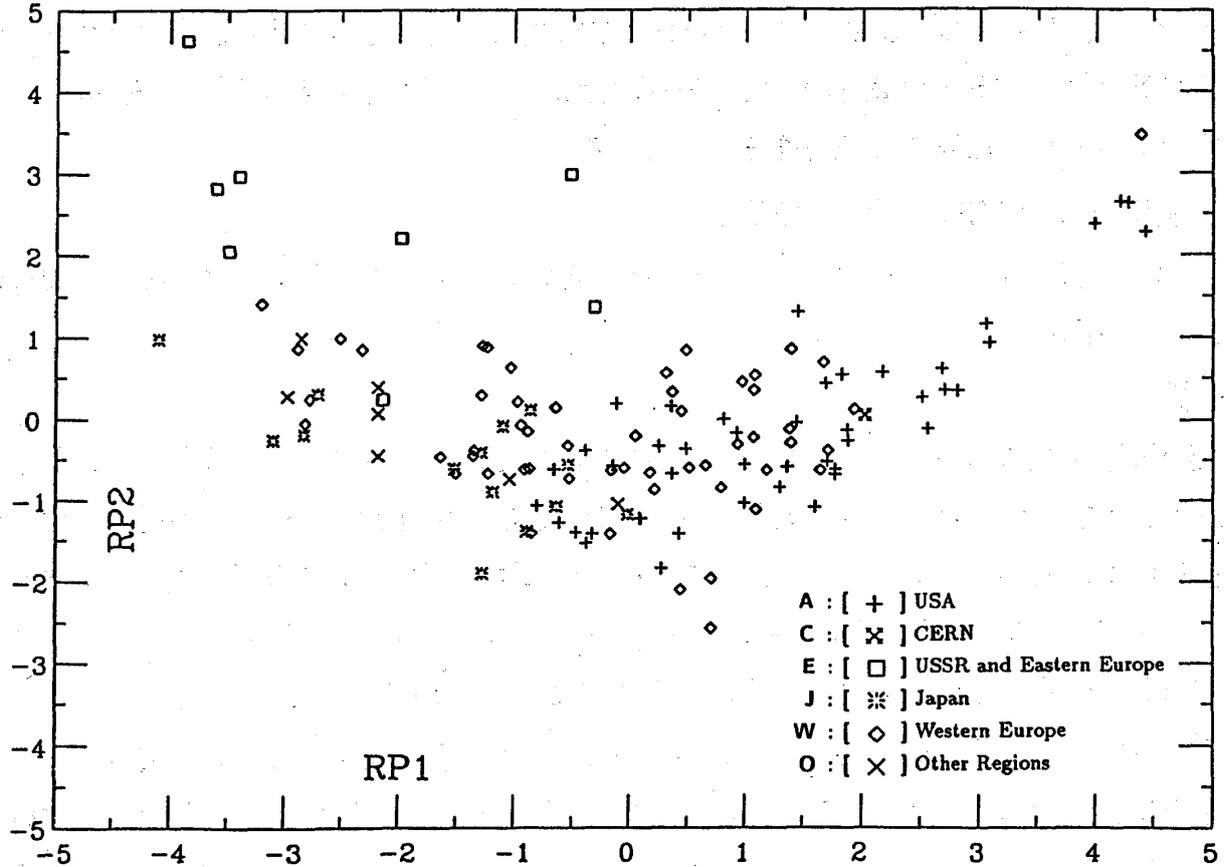
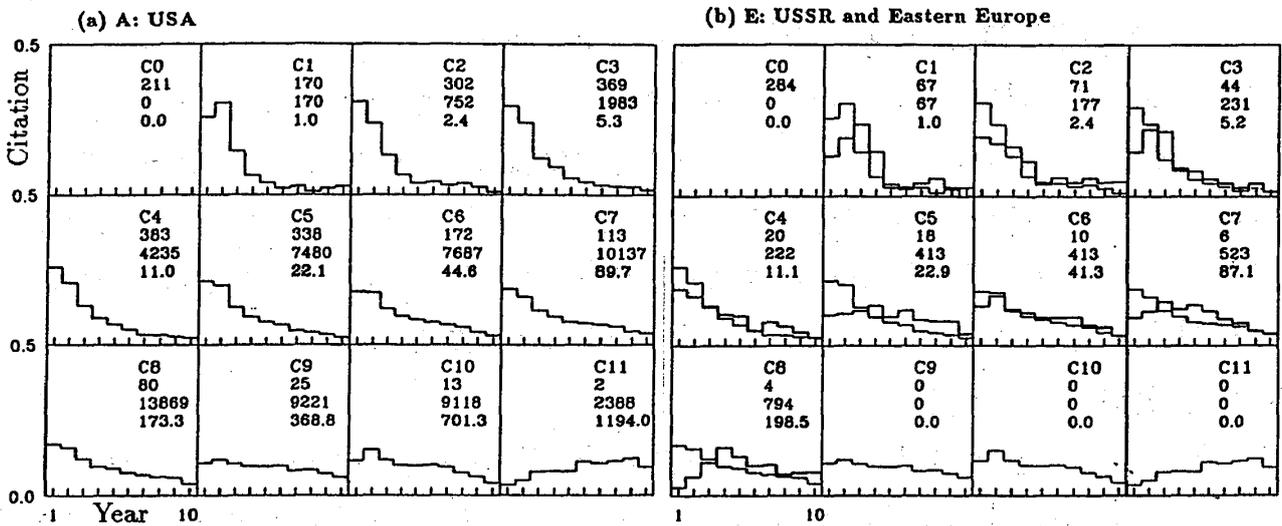


Fig.6



によるクラス分けと論文の寿命がよく相関していることがわかる。Fig.6bはE地域の場合であるが、比較のためにA地域の場合の線も書き入れてある。明らかに、同一のCitationクラスでは、AよりEの方が寿命が長い。言い換えれば、Eの論文はAで書かれていたならもっとciteされたはずである、となる。(日本は定量的にはAとEの間のAよりに位置している。)

以上のデータで見てきたことから、研究情報交換はまだまだ遅れている事、研究者はもっと Active になり得る、あるいは、Active な研究者をもっとふやせる、というのが、物理学研究 Society の現状であろう。私が基研に期待したい事は、この「もっと Active に」のためのグローバルな研究環境を作ることである。

3. 基研の研究情報センター的活動について

これまで述べてきた視点から、基研の研究情報センター的活動を例にとり、具体的なイメージを作る。

a 歴史

基研が設立当初から、「素粒子論研究」の発行など、具体的にこういう面での活動を大きく位置づけてきたことは言うまでもない。しかし、研究情報センターという様な言葉が明確に言われたのは 70 年代に入ってからである。特にプレプリントでの研究成果の公表が主になり、SLAC が新着プレプリントリストのサービスを始めたのが 60 年代末であった。情報の流通が具体的な物の流通として目に見える形になったわけである。

基研では、1976 年頃から準備を始め、DESY で作られ配布されていた文献データベースを京都大学大型センターに移植して国内でのサービスに供した。これを明確な情報センター活動の第 1 期としよう。1983 年には同じく大型センター上で基研プレプリントデータベースの運用を始めたが、その後、国内的にも各研究機関で文献データベース等が整備されてきた。これが第 2 期といえる。1988 年からの第 3 期は、コンピュータネットワークの発展に合わせて、研究情報ネットワークを世界的に作ってきた時期である。

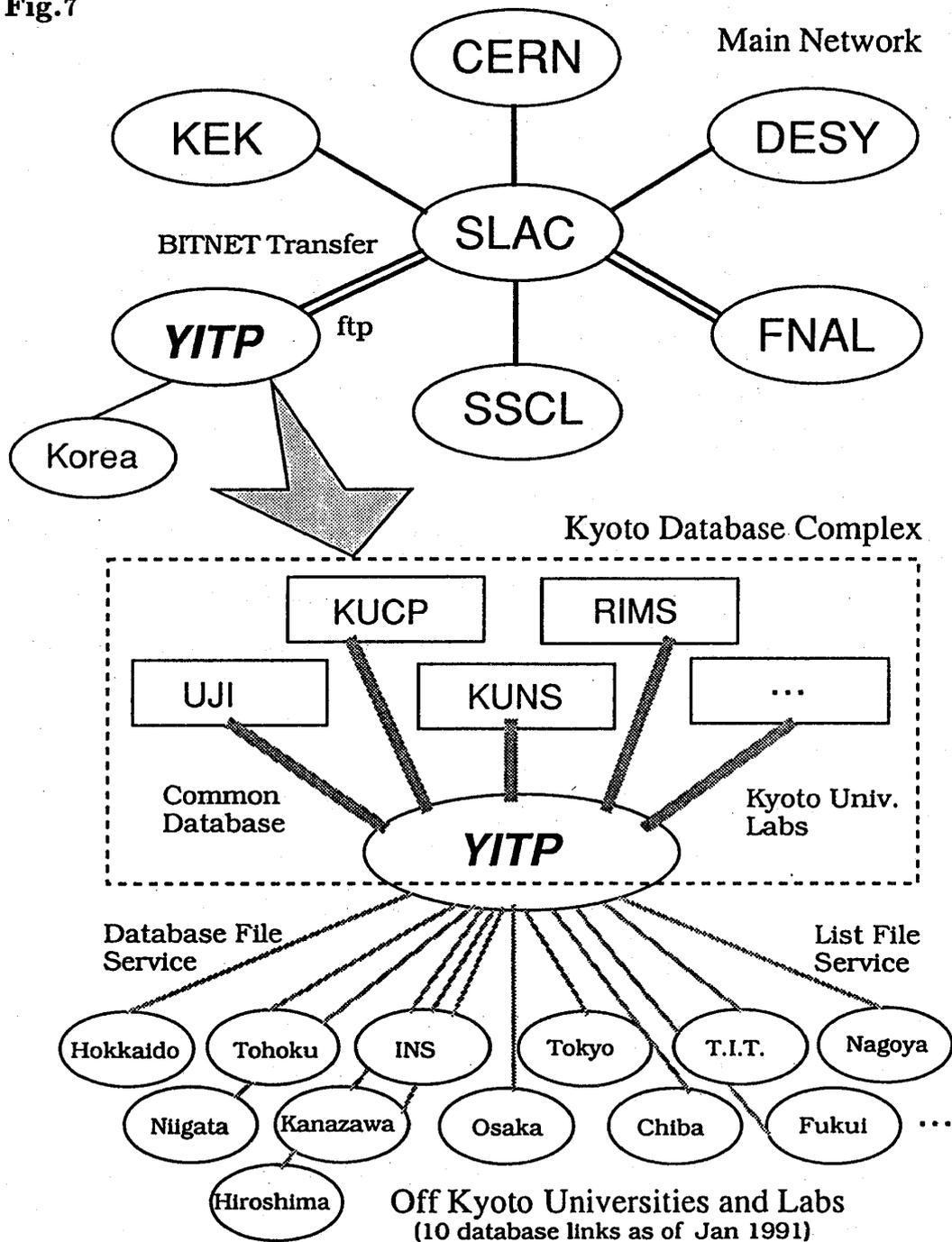
ひとつ確認しておきたいのは、こういう「研究情報センター的なもの」と現実の研究、Activity、国際交流 etc との関係である。つまり、それは、バックグラウンドでありインフラストラクチャーであることは自明であるが、今や必要条件であり、更には、Accelerator であり Driving Force、Initiator であることを以下で強調したい。

b 研究情報ネットワーク

第 3 期で作ってきた研究情報ネットワーク (Fig.7) の特徴は次のようにまとめられる³⁾。まず、Worldwideな研究機関の間のネットワークであり、日常的な共同作業をその基礎にしている。これは、単に少数の研究者が往来するだけの国際交流とは全くレベルが違う、多くの研究機関の間の直接交流である。Fig.7 に見られるネットワークは、階層的な Tree 構造をしているが、これは誰かが決めてこうなったわけではなく、グローバルな構造のSelf-Organizationに他ならない。この構造のなかで、世界の研究機関との国際交流の窓口を担いながら、国内の研究機関の中心となるという基研の姿が自然にできてきた。そし

て、「世界からの窓口」という受け身的な位置ではなく、世界の主要研究機関のネットワークの一翼を基研は担っている。つまり、この研究情報ネットワークに於いては、情報の受信だけでなく、情報の発信も対等に担っている。

Fig.7



この中で、「全国共同利用研としての基研は京都大学の中の関係研究機関を特別扱いすべきではない」という立場からは少し non-trivial な京都大学内の共同関係ができています。この京大内での共同は、京大外の研究機関から見たときにマイナスになるものではなく、むしろ

ろ、全体として非常なプラスになるからこそ、こういう構造ができてきたのである。たとえば、京大内の関係機関では、プレプリントの整理を基研内の世界データベースに完全にリンクして行なっている。以前の部員会議の議論で、プレプリントの取り合的な事が懸念されたが、Fig.8にあるように、旧基研、旧理論研、理学部物理の所有するプレプリントは相補的で、個別に所有する分の約倍の情報が京大全体としては利用できることがわかる。(こういうデータ自身もこの様な共同作業ができて初めてわかることである。) 現在では、基研の北白川と宇治の間では、相手にあるが自分の所に無いプレプリントについて、相手の図書室に電子メールでオーダーすれば、コピーが学内便で自分の図書室に届くシステムが作られている。この様に、京大内の研究機関が協力してプレプリント情報を有効に利用する新しい方法を、研究情報ネットワークの活用によって生み出す事ができたわけである。おそらく、この様な形態を国内さらには世界的に広げて行く事が次のステップへの一つの目標であろう。

もうひとつの新しい特徴は、Koreaとのリンクに象徴されるように、Asia-Pacificのセンター的な役割もまたSelf-Organizeされ始めているということである。これも、「これからはAsia-Pacificだ」と決めて何かを作ってきたわけでは全然なく(実際、Asia-Pacificの基研という概念はよく聞くが、私はまだあまり諸事実を理解していない)、基研が外部からの要求に耳を開いているだけで、自然にできてきたものである。この例は、情報ネットワークが、高尚な上からの議論とは別に(を越えて)、具体的な交流のDriving Forceとなり得る事を示している。

c 次のステップは何か

さて、以上のまとめに基づいて、第4期の研究情報センター的活動をイメージしてみる。まず、いくつかの規定要因をあげる。間違いの無いこととして、空間的距離はなくなる、という事がある。つまり、情報交換についての商業論理がある限り、そして東京までの運賃が市バスより高い限り、東京への市外電話は市内より高いわけである。この事は、地球を覆い尽くすであろうネットワーク構造から階層構造は簡単にはなくなる、という事を意味する。(ネットワークの先進国アメリカでは、回線料金が情報通信量にほとんどよらない動きも出ているが…。)

情報通信手段としてのコンピュータネットワークの優位性は動かない。徒歩から始めて、馬、車、船、飛行機、電信・電話・放送・FAXとなって、その時差は実質ゼロにまでなったが、受け取った情報の再処理の容易性・柔軟性の点でコンピュータネットワークは優れている。階層構造は残るとしても(コンピュータネットワークにおいては、階層構造は管理の面からむしろ必要とも言える)、コンピュータの(インフラでの)スピード・容量は、おそらく21世紀には今日の1000倍以上になると考えられる。こういう面でのアメリカと比

した時の日本の遅れは大きい。グローバルな研究環境という視点からは、ウルトラのスピード・容量 (日本は既にトップ) よりもインフラ側が重要である。

そして結局、コンピュータネットワークは21世紀を特徴付ける文明環境であろう。理論物理学の研究環境は常に最先端の文明環境を要求し、活用というより開発してきたのが歴史であった。例えば、高エネルギー関係の文献データベースネットワークのデータベースシステムはスタンフォード大学から供給されているSPIRES(Stanford Public Information REtrieval System)であり、このデータベースシステムは構造を持ったデータを効率よく管理する優れたシステムとして、図書館の管理等多くの大学で採用されている。実は、このSPIRESの名前は最初は、Stanford Physics...であった。つまり、SLACを中心とする物理関係の人々が物理研究の必要に迫られて一から開発したものが今やPublicに使われている。もちろん、研究環境の構築が文明環境を先導する事は、物理に限られた事ではない。最近目についた2つの文を引いておきたい。ひとつは、JUNETの創始者である慶応大学の村井純氏⁴⁾:「学術コミュニティのコンピュータネットワークは未来の人類の総合的なコミュニケーションの原点」、もうひとつは、京大名誉教授(人文研)の樋口謹一氏⁵⁾:「インターナショナルからトランスナショナルへ」。これらは、今後基研が作るべきグローバルな研究環境の性格についての貴重な示唆をしているとわたしには思える。

以上をまとめると、第4期といえる基研情報センター的活動のイメージは、グローバルな環境をトランスナショナルに創造し発展させる事、研究情報環境という面での、グローバルなベストコンディションの整備に寄与する事、となろう。具体的に少し例をあげると、前でも述べたが、Document Server的なもの(文献の中身全体の情報をネットワーク上に持ち、即時に印刷出力を得ることができる。既に世界的に議論が始まっている。)がひとつのステップになる。また、情報発信手段としては論文雑誌が一番重要であり、基研が深く関わってきた日本の雑誌Progressについて、情報ネットワークという環境の下での大幅な衣替えを期待したい。

4. Operating System “基研” Version 2

今、新しいスケールの下での基研の将来像を議論しているわけだが、これはいわば、基研というOperating SystemのVersion 2の仕様を決めるという作業に擬すことができる。Operating System(OS)というのはコンピュータ用語であるが(シンポジウム当日、私の説明不足のため甚だしく誤解をされた方があった様で、お詫びいたします。)、単純な機能のハードウェアの上に作られた、色々なソフトウェアが走れるプラットフォームを指す。例えば、パソコンのOSはMS-DOSと呼ばれ、人間が情報をファイルとして扱い、そ

れを複製する事が簡単にできるコマンド (COPY コマンド) を用意してくれている。ここでいう“基研”というのは、建物の事でも組織の事でもなく、研究者によって議論され運営されているその内容そのものを抽象的に意味している。従って、文部省の一組織であり、その規則と予算使用の部分がコンピュータでいえばハードウェアに対応し、その上に研究者が自分たちの理想を持って色々な共同利用、国際交流、人事等のシステム、即ち Operating System を作ってきたわけである。

コンピュータの世界での Operating System も大きく進化してきた。端的に言えば、独裁的 CPU に多数の非知性端末がぶら下がるという構造に対応した古い OS から、未来の姿として、ネットワーク、サーバー(データベース)、分散環境等をキーワードとする OS が追求されてきた。この未来型の OS のプロトタイプとして最近の流行が UNIX である。私のいう結論のフレーズ、「Operating System “基研” Version 2 は UNIX で行こう」はまさにこの意味で UNIX を引いている。UNIX がコンピュータ会社ではなく研究者によって作られてきた事も上の比喩の整合性を補強する事実である。

私は、OS “基研” Version 1 を上の意味での古い OS に対応させてはいない。むしろ、理想の新しい OS に対応する物を極めて不十分なハードウェア資源の上で実現しようと苦闘して来たのが“基研” Version 1 であったと考えている。コンピュータの OS の場合は OS の上で別の OS を動かす事が良くある。日本には「科研費」という巨大な OS が存在しているが、素粒子論グループ等は、この上に別の OS を動かそうとしていると見れる。

これらのアナロジーは単なる言葉の対応以上の意味を持っている。なぜなら、良い Operating System の目標は、限られたハードウェア資源を如何に有効に使って、人間が如何に快適に目的を果たせるかということにあり、コンピュータの問題として研究されてきた概念の整理等がそのまま“基研”という Operating System を考える場合にも助けになるからである。更に、先に述べたように、研究環境どころか文明環境としてのコンピュータネットワークが地球を覆い始めている。ということは、Operating System “基研” の機能の実現を計る実際の (物質的) 環境基盤が実はコンピュータネットワークであり、仮想的に導入した OS の実現が実は本当のコンピュータの OS によって実現されていくわけである。この事は、例えば、物理的な構造としてのコンピュータネットワーク (の結線の仕方) と、抽象的な意味での“基研”の作る研究環境の構造が一致する事に現われる。一致するだけなら、それだけの結果論であるが、それぞれがダイナミックに創造されていく過程では、それらが共鳴し、それぞれが他方に対する Driving Force となり得る。この共鳴が、Operating System としての“基研”という概念を単なるアナロジーを越えたものにしていく、と私は実感している。

というわけで、私の描く 21 世紀の基研の将来像は、Fig.9 の様になる (Yukawa Inst. の丸...)。一番外側の丸が地球、階層構造は残りながら、点を線で結ぶのではない面的な構造 (トランスナショナル) があり、基研はそれをグローバルな研究環境として位置づけて生み出し支えているが、誰も日常的には意識しない。そして、「全国共同利用」とか「国際交流センター」とかいう言葉自身がなんとなくもどかしく (あまりにも自明だから ...) 既に昔の言葉になっている...

Fig.8

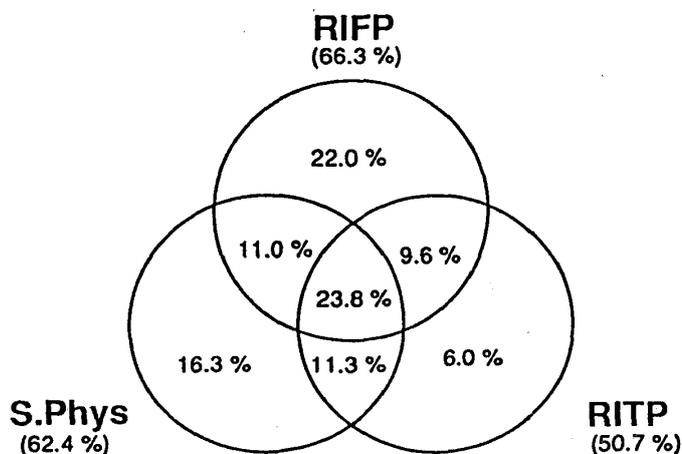
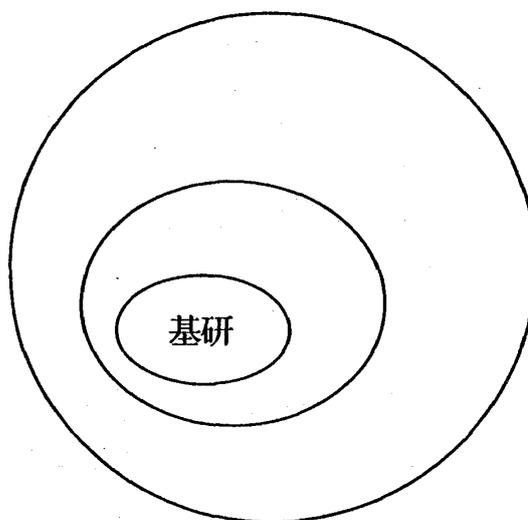


Fig.9



Preprint Overlap, Mar 1990.

文献

1. "Research Activity in Theoretical High Energy Physics", K-I. Aoki, M. Bando, M. Toya, RIFP-702(1987)1-71.
2. 「目で見える素粒子論のうつりかわり」、青木健一、登谷美穂子、坂東昌子、Parity 3-7(1988)75-80; 「データベースによる研究アクティビティの分析 — 高エネルギー物理学(1974-1987)の場合 —」、青木健一、登谷美穂子、『京都大学大型計算機センター広報』、21-4 (1988) 239-245.
3. 「コンピュータネットワークと研究情報」、青木健一、『科学』 60-2 (1990) 67-68.
4. 「学術コミュニティのコンピュータネットワーク」、村井 純、『科学』 60-11 (1990) 725-725.
5. 「共同研究 — 過去・現在・未来 —」、樋口 謹一、『京大広報』、397 (1990) 993-993.