Spatial Analysis Visualized

How many words are a graphic worth?

Some examples from my work.

An introduction using eleven examples.

W. Tobler
Geographer

Abstract:

Some quantitative and computational procedures are presented in simple graphical form.

A few examples from my research in geography. Included are spatial and temporal lag effects, comparison of movement patterns, consequences of changing spatial resolution, migration potential and fields, impact of biproportional adjustment on movement tables, coalescent cities, explanation of multidimensional scaling iterations, the transform-solve-invert paradigm, a map comparison method, mass preserving reallocation of spatial data, and potentials from asymmetry.

First Visualization

Spatial and temporal lags in movement tables illustrated and made obvious.

Nine Region Migration Table

US Census 1973

	1	2	3	4	5	6	7	8	9
1 New England	—	180,048	79,223	26,887	198,144	17,995	35,563	30,528	110,792
2 Mid-Atlantic	283,049	—	300,345	67,280	718,673	55,094	93,434	87,987	268,458
3 East North Central	87,267	237,229	—	281,791	551,483	230,788	178,517	172,711	394,481
4 West North Central	28,977	60,681	286,580	—	143,860	49,892	185,618	181,868	274,629
5 South Atlantic	130,830	382,565	346,407	92,308	—	252,189	192,223	89,389	279,739
6 East South Central	21,434	53,772	287,340	49,828	316,650	—	141,679	27,409	87,938
7 West South Central	30,287	64,645	161,645	144,980	199,466	121,366	—	134,229	289,880
8 Mountain	21,450	43,749	97,808	113,683	89,806	25,574	158,006	—	437,255
9 Pacific	72,114	133,122	229,764	165,405	266,305	66,324	252,039	342,948	—

This is an example of a census migration table. There are also (50 by 50) state tables and (3100 by 3100) county by county tables.

Can you comprehend a table of over 9 million numbers?

I know that can't!

It is well known that spatial autocorrelation is present in migration tables.

See, for example;

Curry, 1972, “A spatial analysis of gravity flows”, Regional Studies, 6: 131-147

Griffith & Jones, 1980, “Explorations into the relationship between spatial structure and spatial interaction”, Environment and Planning, A,12:187-201

Accounting for these effects in a model is difficult.

See, for example:

LeSage & Pace, 2007, ”Spatial Econometric Modeling of Origin-Destinations Flows”, http://ssrn.com/abstract=924609

There is a great deal of spatial coherence in the migration pattern

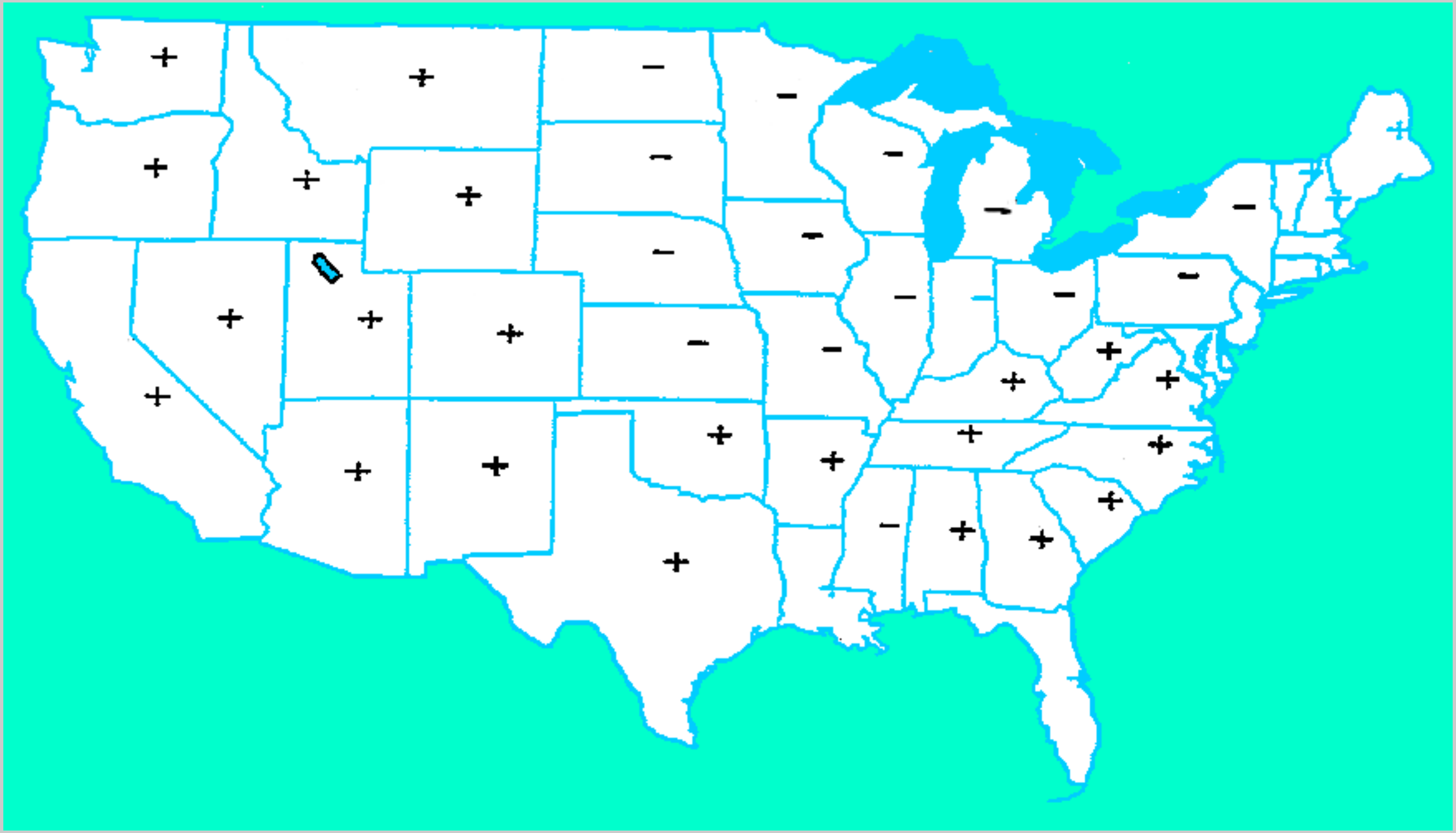
In the US case the state boundaries hide the effect.

Therefore they should be omitted.

There is also temporal coherence.

W. Tobler, 1995, “Migration: Ravenstein, Thornthwaite, and Beyond”, *Urban Geography,* 16(4): 327-343.

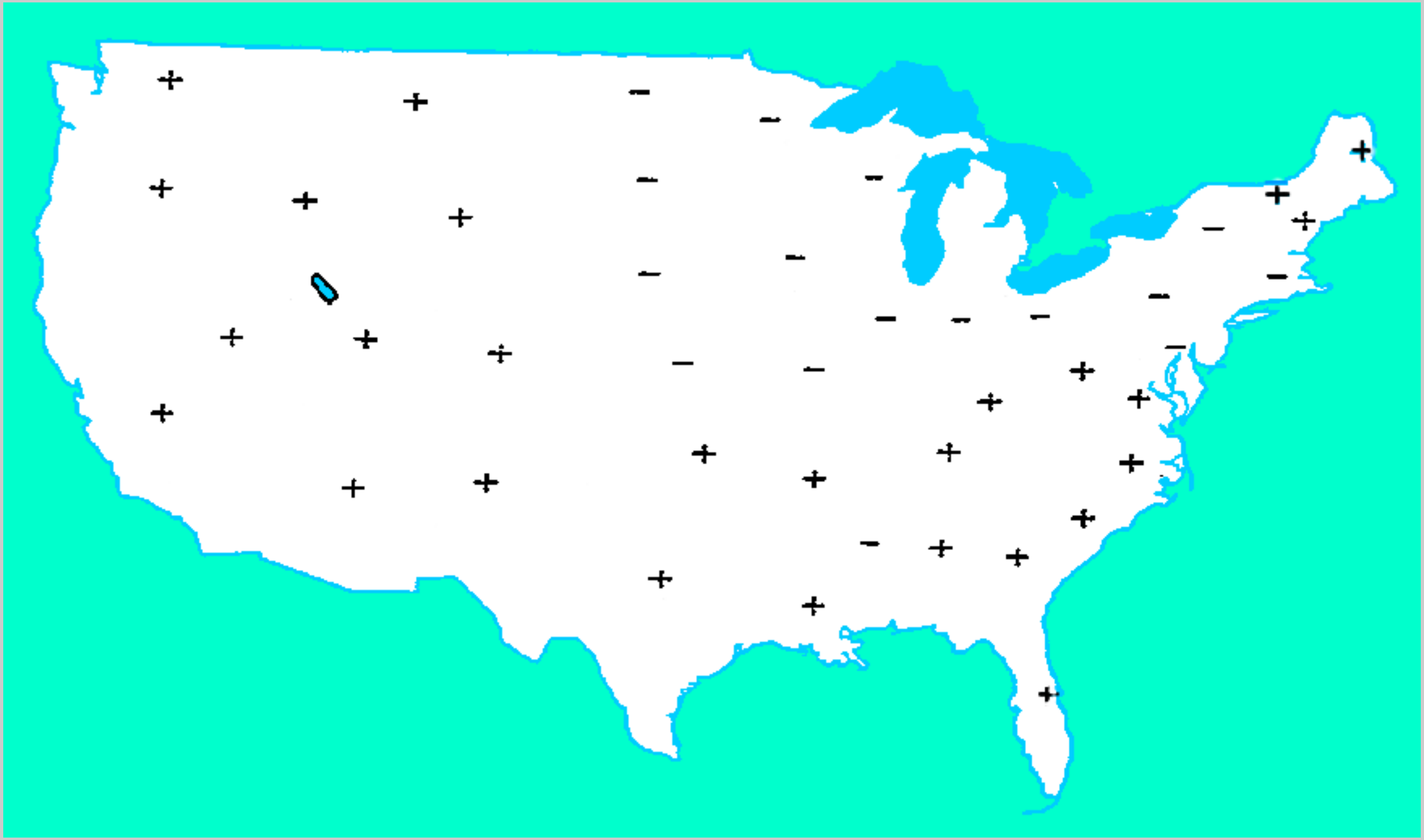
The Population Change Information Can Be Positioned Locationally using centroids



This type of information is often shown using choropleth maps. I consider this method to be simpler and superior.

Observe the spatial autocorrelation and how this is brought out more clearly by omitting the collection unit boundaries, as on the next map.

Population Change at State Centroids

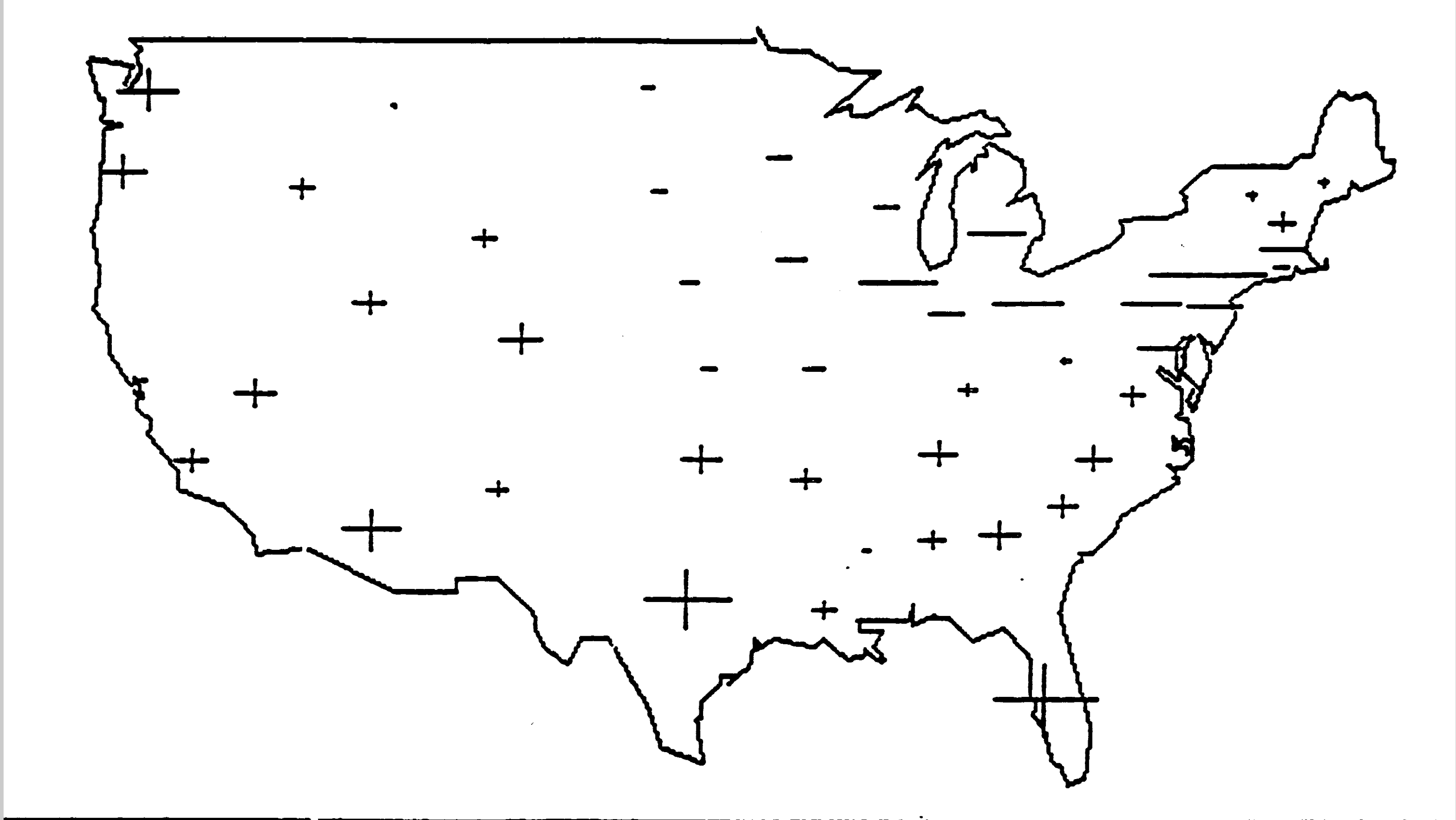


The map is even better if the symbol size is made proportional to the magnitude of the change, as on the next map

Gaining and Losing Migration States

Symbol positioned at the state centroids, and proportional to magnitude of the change.

The map is based on the marginals of a 48 x 48 state to state migration table and shows the accumulation and depletion places. Draw a boundary around the losing states. This demonstrates that states are not the appropriate size for studies of migration and also that there is a great deal of correlation amidst state migration data.



The previous maps make it clear that there is spatial autocorrelation.

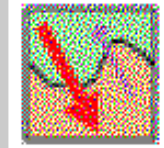
It is not necessary to invoke complicated statistics to see this.

The maps make it obvious.

That there is temporal coherence
is also easily visualized using movement maps.

CSISS.org/Spatial Tools/Tobler's Flow Mapper

Tobler's Flow Mapper



Background

I claim that geographical movement is of crucial importance. This is because most change in the world is due to movement; the movement of people, ideas, animals, insects, disease, money, energy, or material. One way of depicting and analyzing geographical movement is by way of geographical maps. A convenient and rapid method of displaying movement data on such maps is therefore very useful. A flow mapping program is one approach to this objective. For in depth information see csiss.org/Spatial Tools:

- Flow Mapper program and tutorial; and my web site for several publications on migration.

About Flow Mapper

In 2003 CSISS supported a short effort to produce an interactive flow mapping program. The result is a new Windows-based version of a 1987 program by Waldo Tobler. This original application has been updated by David Jones using Microsoft Visual Basic.Net and Scaleable Vector Graphics for map rendering. It requires as input locational coordinates and information on the interaction between the places. Additional input may include place names and a file of boundary coordinates (for a background map). The user has several menu options for producing a map. The program allows for the production of a total movement maps shown by volume-scaled bands, net movement given by scaled arrows, or simultaneous two-way moves.

Some nice properties of the program

- Simple and quick flow map preparation - **GIS Not Needed!**
- Extensive color styles available. Black & white too.
- Hovering over a band or arrow gives the magnitude.
- Hovering over a centroid gives its label.
- Two-way, total, or net movement maps.
- Many to many, one to many, or many to one maps.
- Easy threshold choice. Some statistics made available.
- Size dependent only on memory availability.
- Multiple output formats.
- Non-geographic flows within firms, industries, organizations, too.
- Help file included.
- Microsoft Windows compatible.

A sequence of migration maps

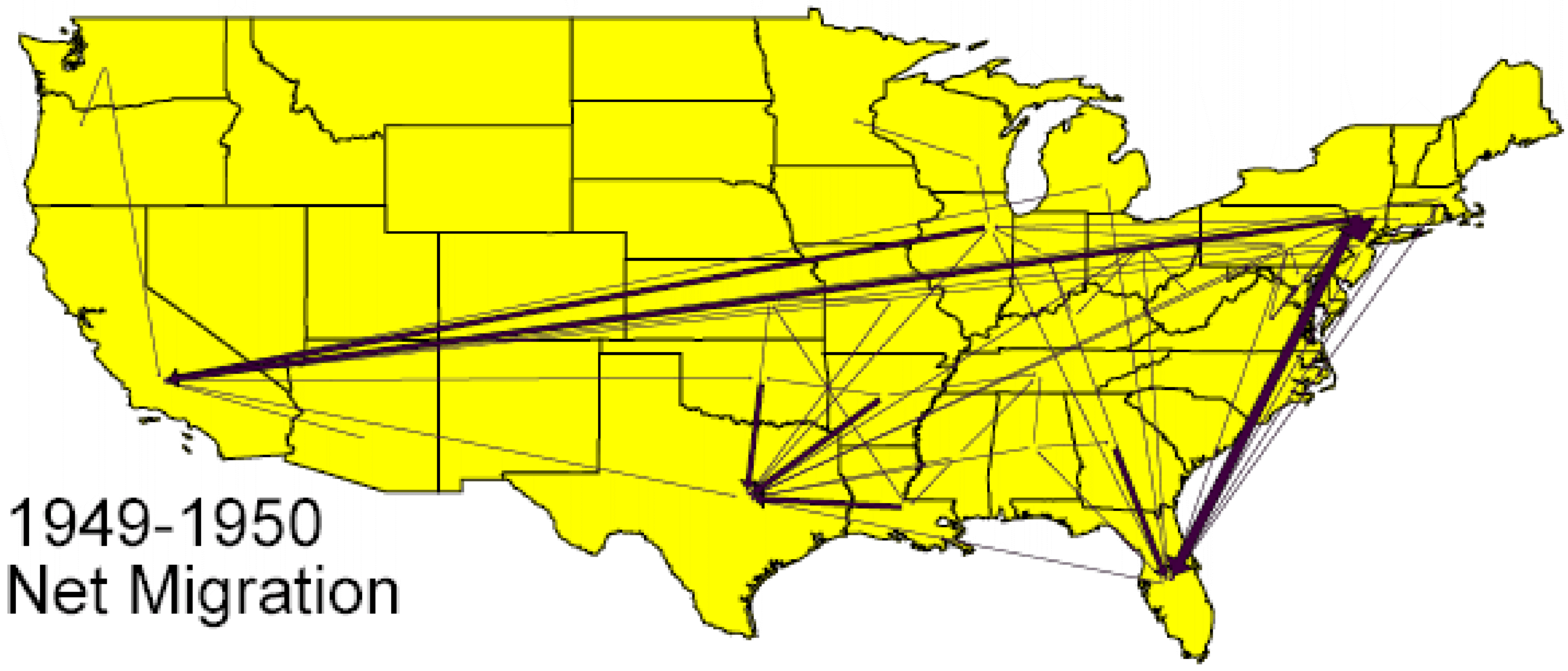
Based on US Census Bureau data

The dust bowl time

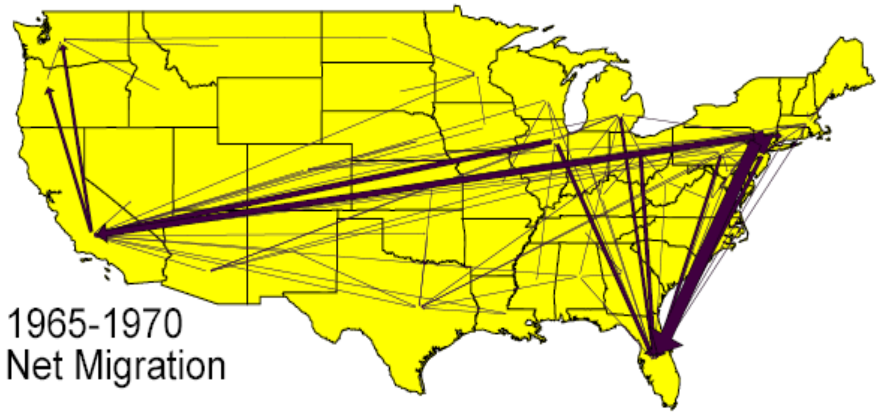
Only one year of data

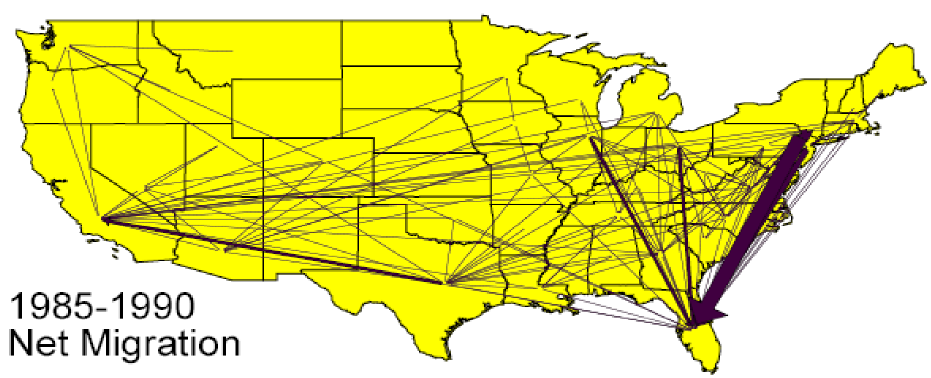
In the other cases there are five years of data

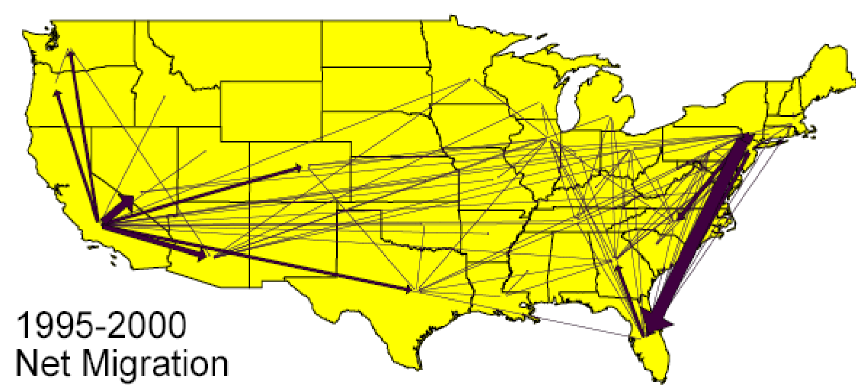
Florida shows up



Four decades of migration maps







Second Visualization

Comparison of spatial movement patterns

Movement of different groups.

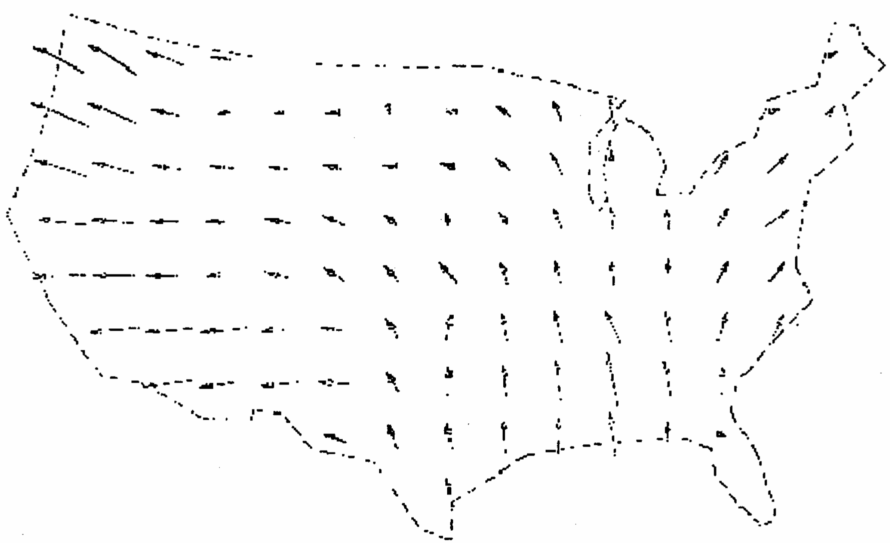
Movement at different times.

Maps produced using a mathematical migration model.

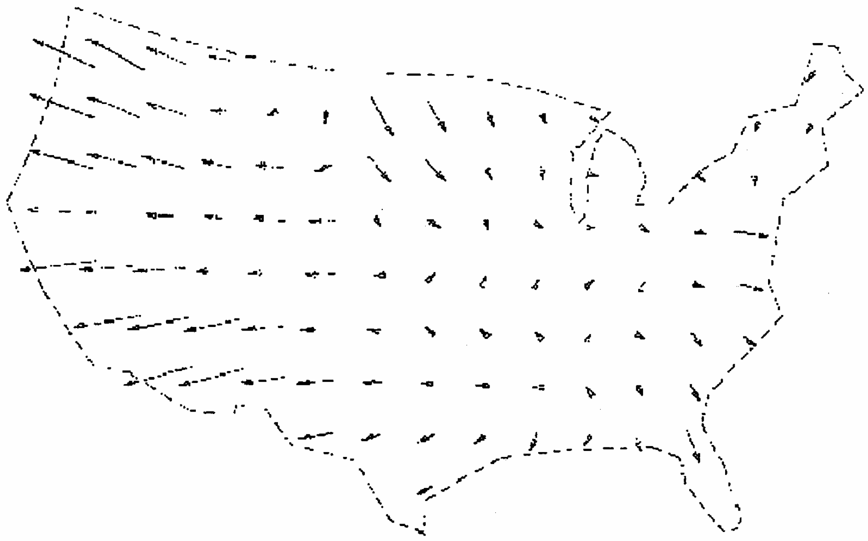
Comparison of non-white and white net migration

1935-1940 five year census migration data, by state

Non White migration



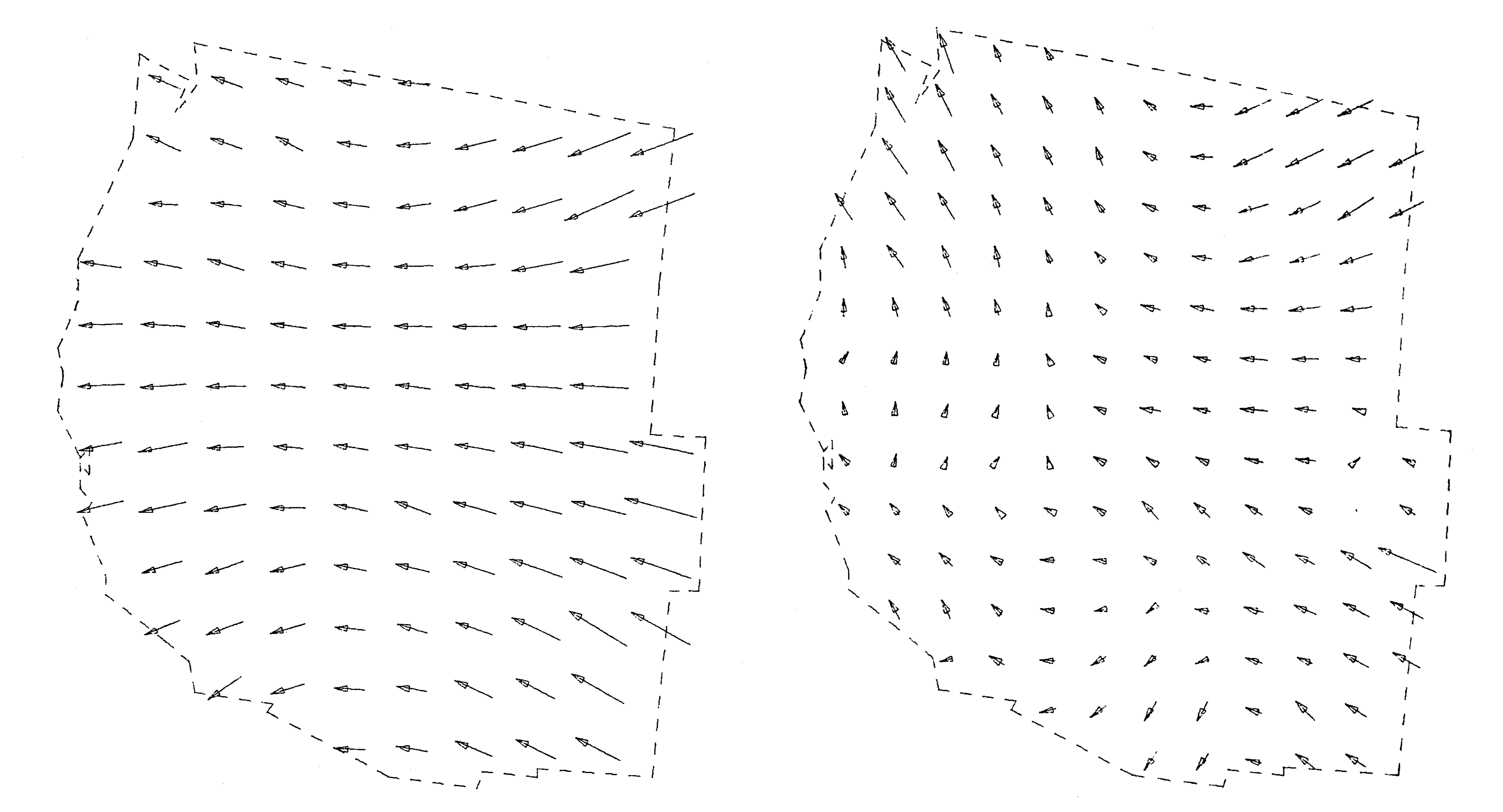
White migration



Comparison of net migration patterns at an interval of thirty years.

Migration in the Western United States by State Economic Areas

Left: 1935-1940. Right: 1965-1970.



Third Visualization

The effect of resolution on migration data.

Social data are often made available in a hierarchy of administrative units.

Moving through the hierarchy changes the resolution and this acts as a spatial filter.

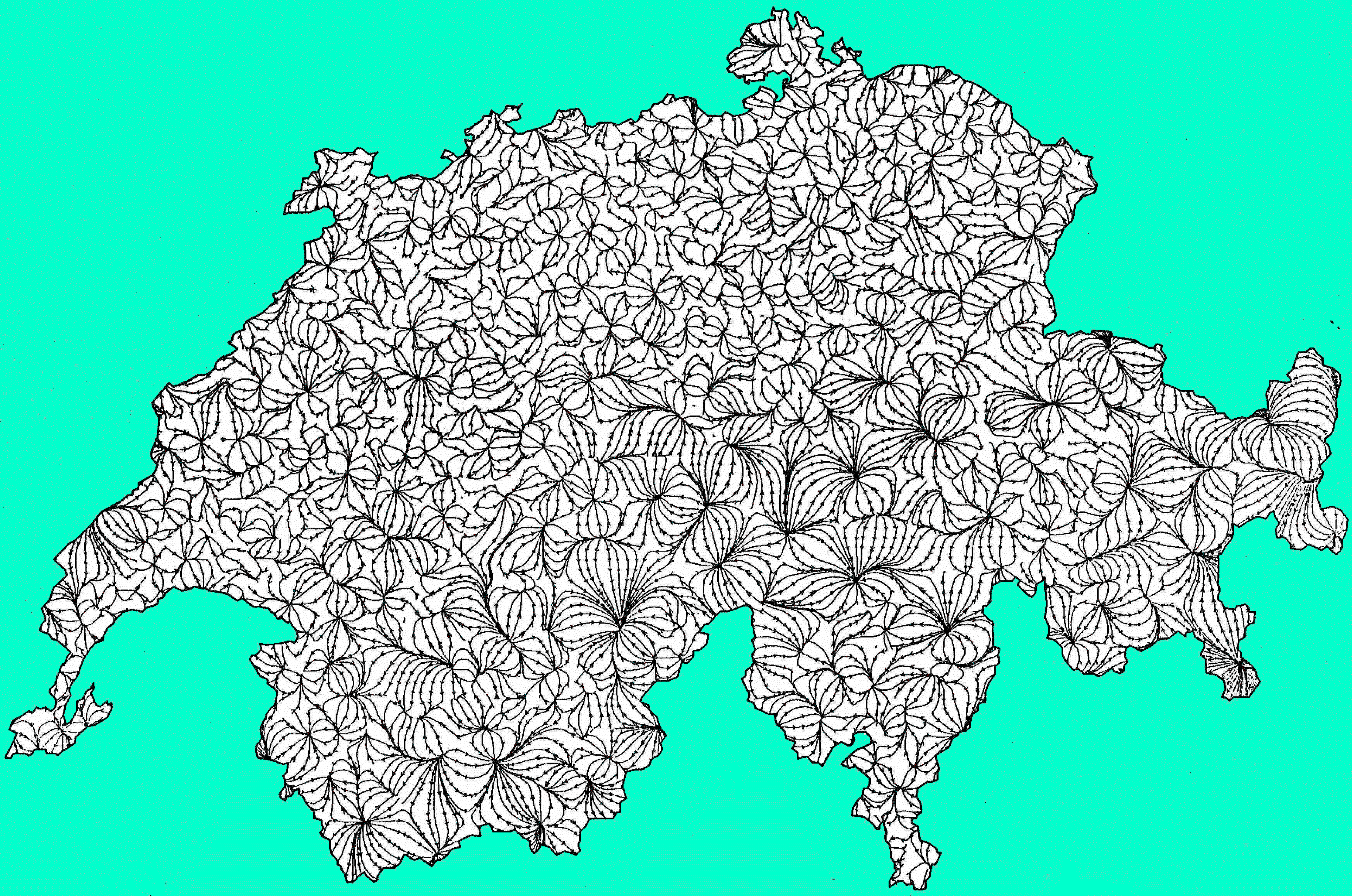
This is shown by migration vector fields at several levels of resolution for Switzerland.

41,293 sq. km.   Average resolution = sq. root (Area / Units)

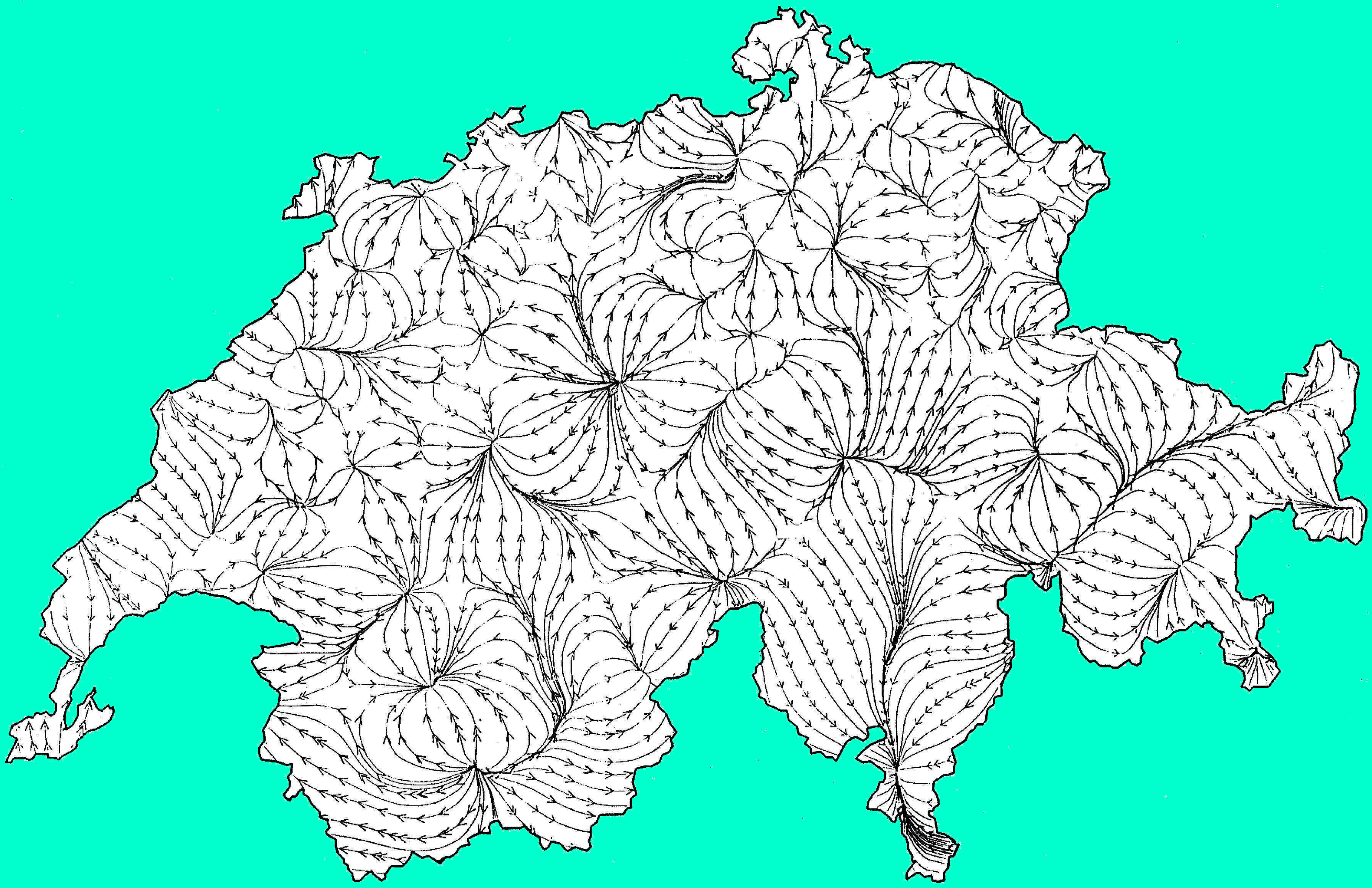
3.6 km resolution (3090 Gemeinde)
14.7 km resolution (184 Bezirke)
39.2 km resolution (26 Kantone)

Maps by Guido Dorigo, University of Zürich, based on a program by Waldo Tobler

Migration “Turbulence” in the Alps. 3090 units - 3.6 km resolution

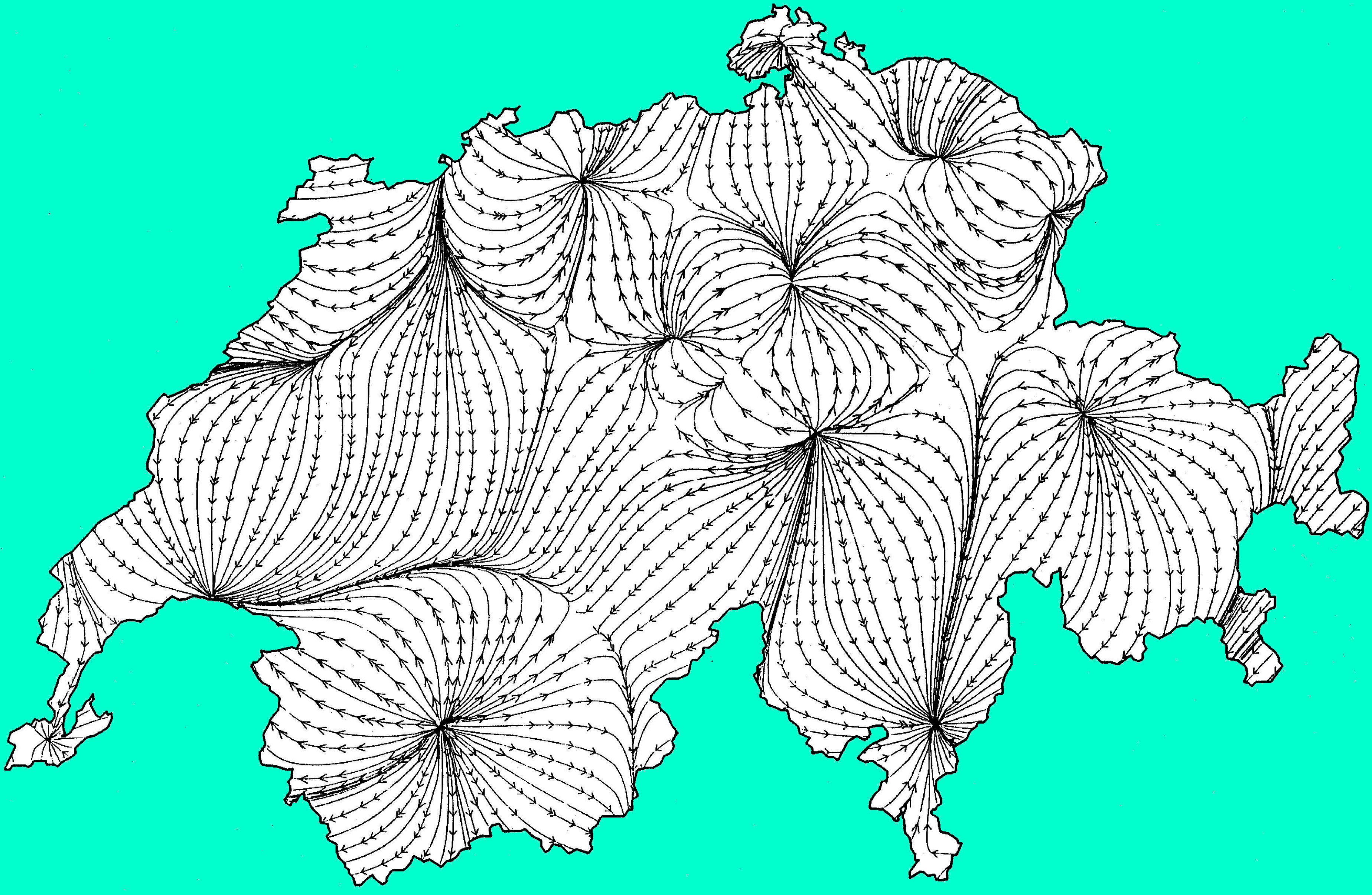


Less of the Fine Detail. 184 units - 14.7 km resolution



The Broad Pattern Only 26 units - 39.2 km resolution

Changing the resolution has the effect of a spatial filter.

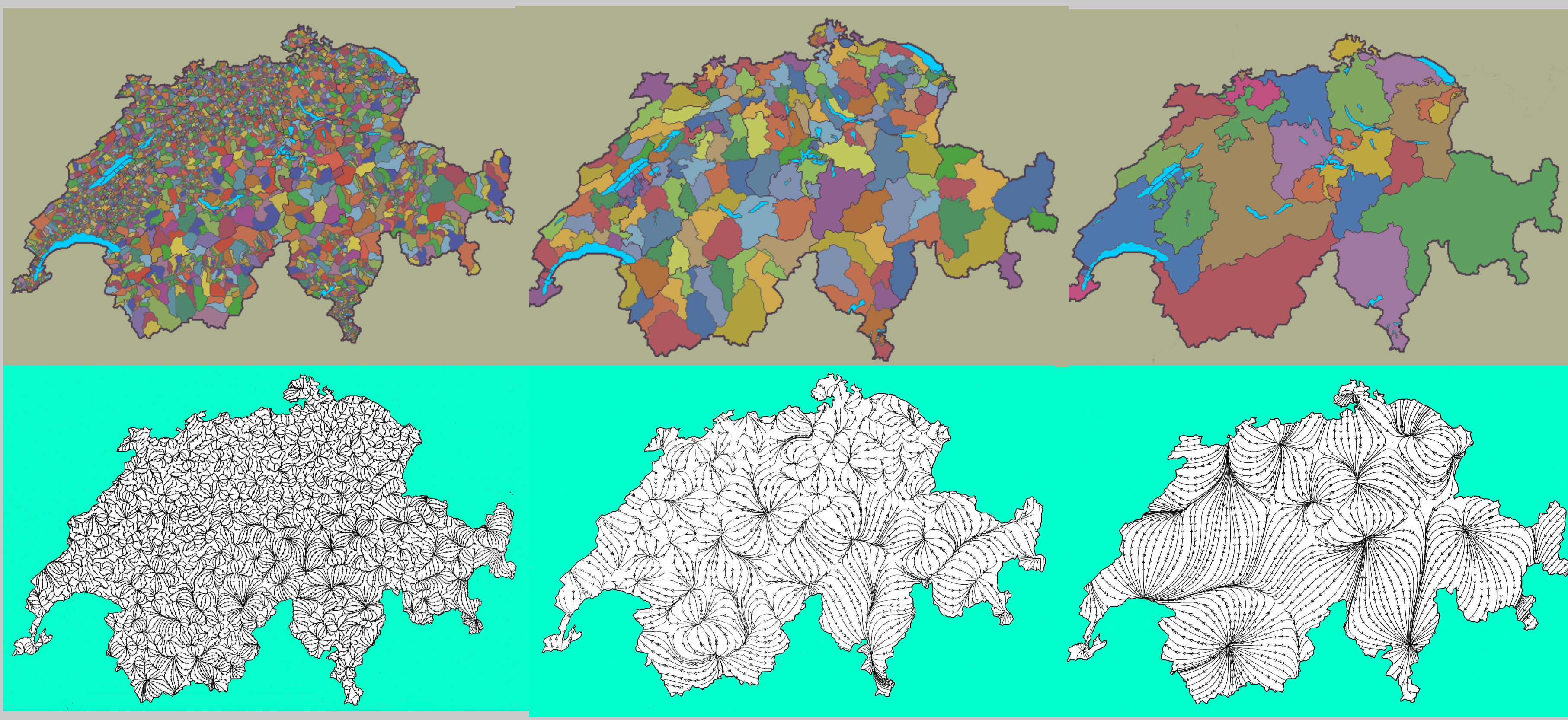


Three levels of administrative units and three levels of migration resolution all at once.

Notice that the resolution is not uniform thruout the country.

On Spatial Filtering see: J. Holloway, 1958, "Smoothing and filtering of time series and space fields" Advances in Geophysics, 4: 351-389.

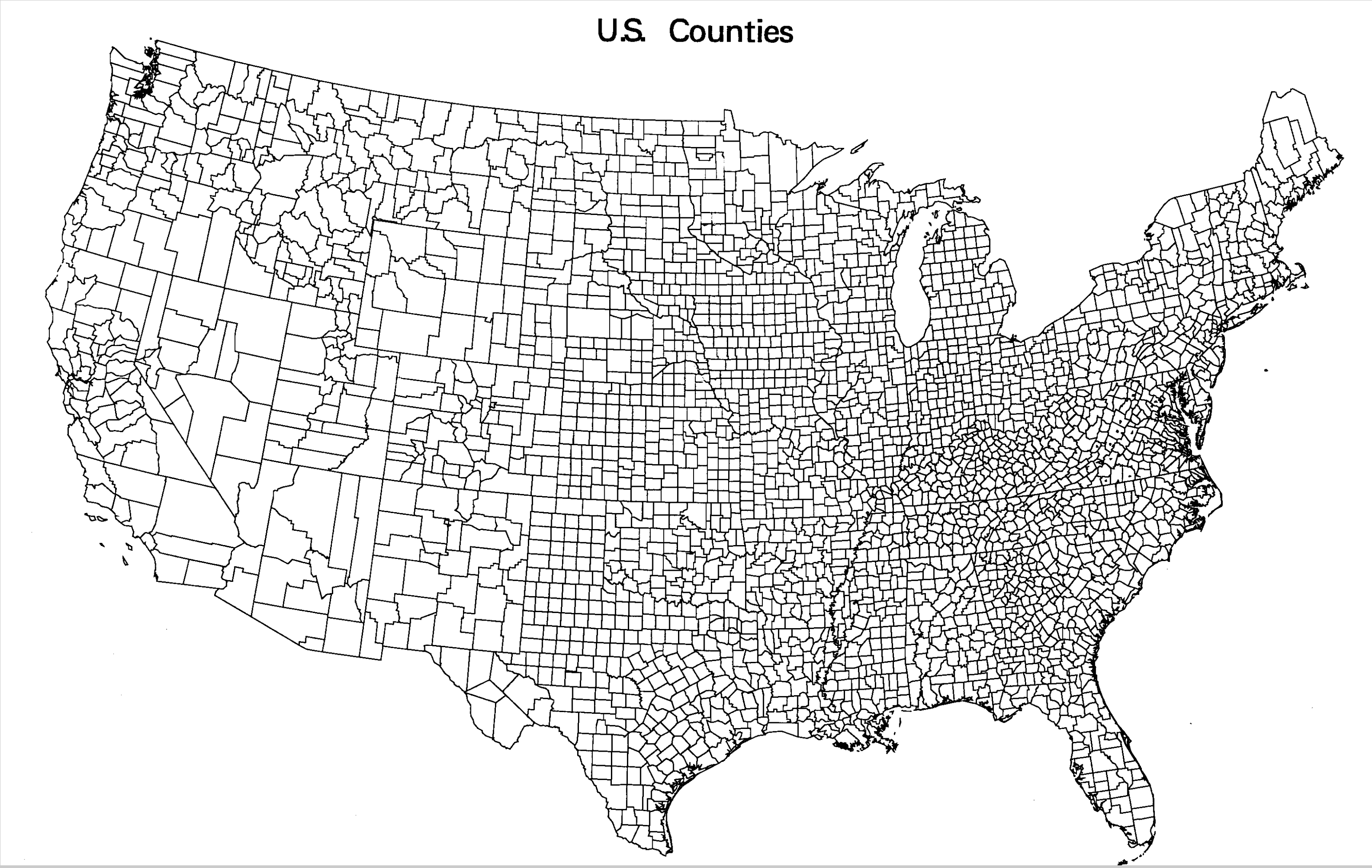
Communities	Districts	Cantons
3090 units 3.6 km	184 units 14.7 km	29 units 39.2 km



County Units

3100+ units

If you got this kind of resolution in photographic film you would reject it, wouldn't you?



Average resolution ~55 km. Patterns >110 km detectable.

Still not sufficient to see movement within cities.

Fourth Visualization

Displaying the results of a migration model.

The model is given as a pair of partial differential equations for a uniform fine mesh covering the region, namely

$$\partial^2\mathbf{R}/\partial u^2 + \partial^2\mathbf{R}/\partial v^2 = \mathbf{I}(u,v) - 4\mathbf{T}(u,v),$$

$$\partial^2\mathbf{E}/\partial u^2 + \partial^2\mathbf{E}/\partial v^2 = \mathbf{O}(u,v) - 4\mathbf{T}(u,v),$$

assuming that the **E** ('pulls' ☺) and **R** ('pushes' ☹) are differentiable spatial functions and that **I**n-movement, **O**ut-movement, and **T**urnover are continuous densities given as functions of the Cartesian coordinates u and v.

This is the model used for the preceeding slides.

G. Dorigo, W Tobler, 1983, "*Push-Pull Migration Laws*", Annals, Association of American Geographers, 73(1):1-17.

In this continuous two dimensional migration model we have a coupled system of two simultaneous partial differential equations

These equations can be combined to yield either gross movements or net movements.

For the simultaneous movement in both directions at each pair of plaes **add** the two equations to get the 'turnover' potential, **T**.

The result is Helmholtz's equation.

One of the "Laws of Migration" is that in and out movements are nearly the same. Ravenstein 1885.

For the net movement we need only the difference between the ‘in’ and ‘out’ at each node for the ‘attractivity’ potential, as follows:

By **subtraction** from the previous equations, we have

$$\partial^2 \mathbf{A}/\partial u^2 + \partial^2 \mathbf{A}/\partial v^2 = \mathbf{I}(u,v) - \mathbf{O}(u,v),$$

where **A** = **E** - **R** (‘pull’ ☺ minus ‘push’ ☹) can be thought of as the attractivity of each location.

This is the well-known Poisson equation for which numerical solutions are easily obtained.

The right hand side, **I**n minus **O**ut, is the amount of change.

Once **A**(u,v) - the potential - has been found from this equation, the net movement pattern is given by the vector field,

$$\nabla = \text{grad } \mathbf{A},$$

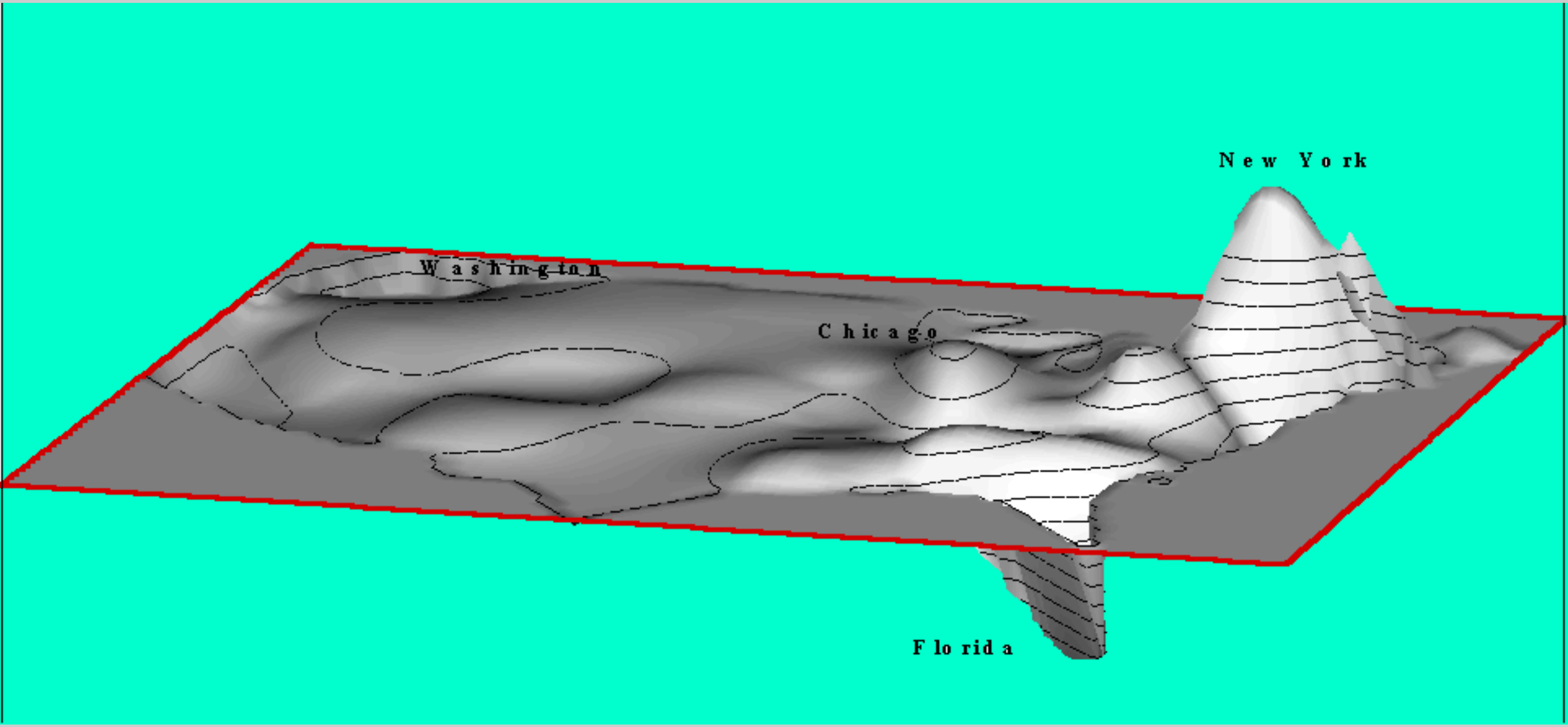
or by the difference in potential between each pair of mesh nodes.

Solving the Poisson equation for the potential gives

The Pressure to Move in the US

Based on the continuous spatial quadratic model

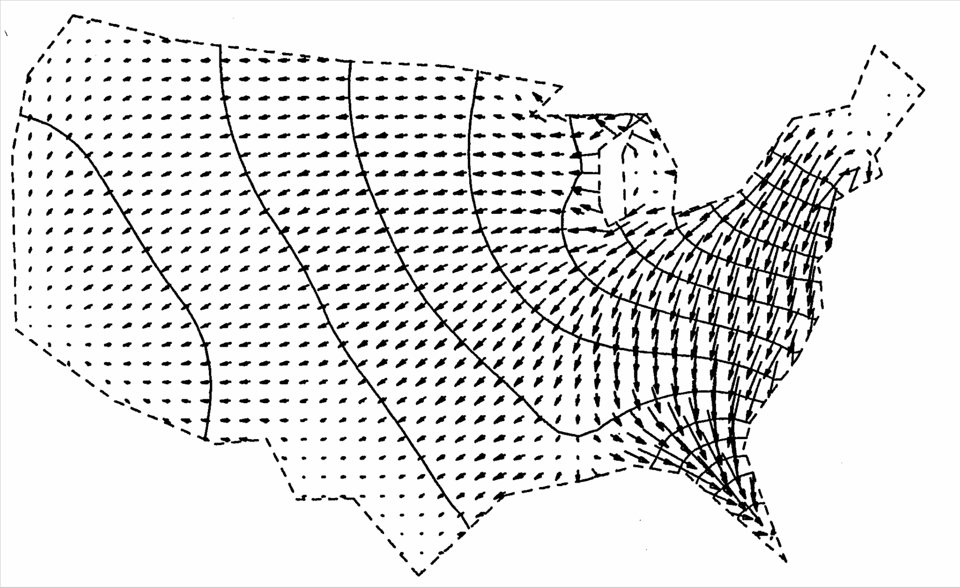
Movement is from high to low; using state data

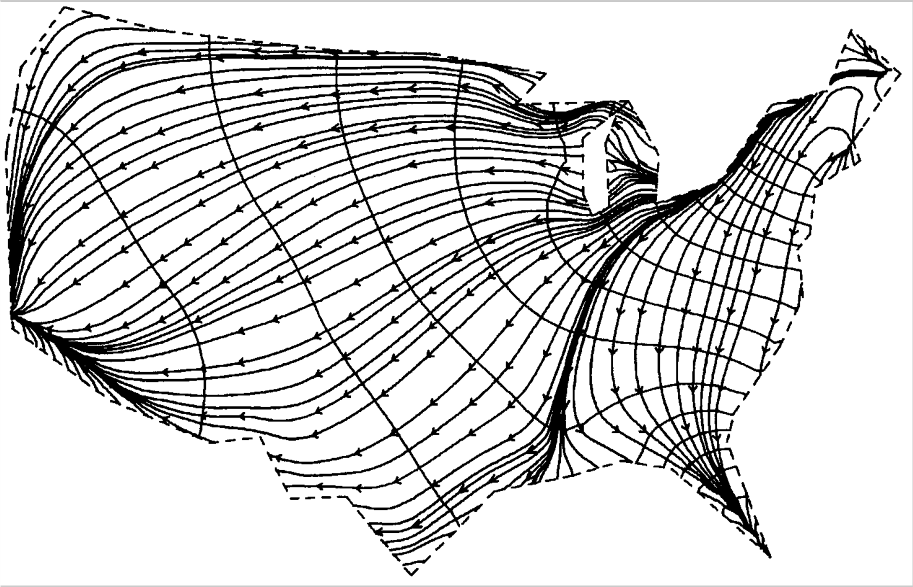


An alternate view of the same result

The migration potentials shown as contours

and with gradient vectors connected to give streaklines





Another Example

In the United States the currency indicates where it was issued

For bills this is the Federal Reserve District.

Coins contain a mint abbreviation.

You can check your wallet to estimate your interaction with the rest of the country.

For Europe use the Euro coins.

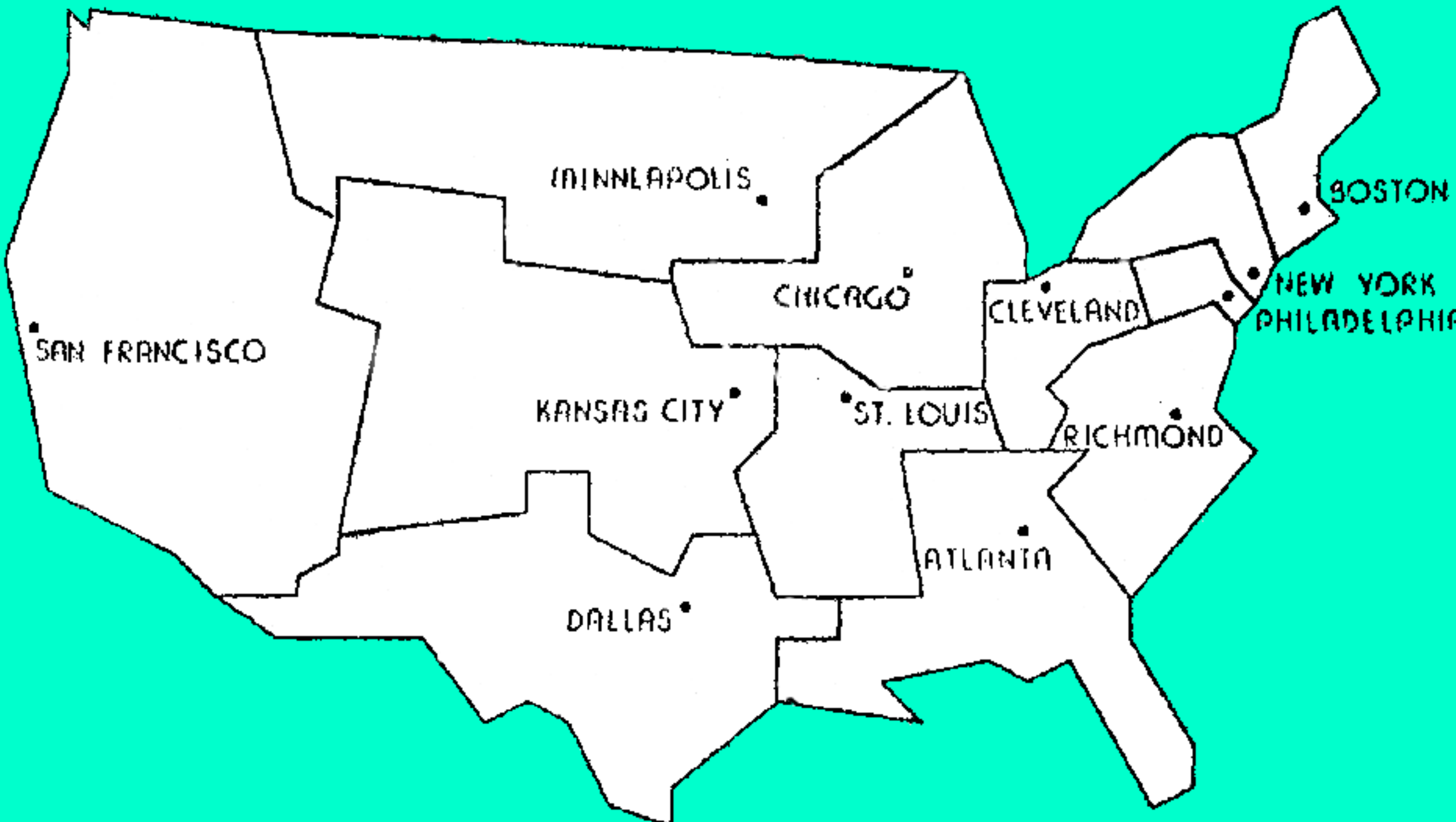
Dollar Bill
(Federal Reserve Note)

Issued by the 8th (St. Louis) Federal Reserve District.

(H is the 8th letter of the alphabet)

The 12 Federal Reserve Districts

(Alaska and Hawaii omitted)



The Table of Dollar Bill Movements

was obtained from MacDonalds outlets throughout the United States.

Source: S. Pignatello, 1977, *Mathematical Modeling for Management of the Quality of Circulating Currency*, Federal Reserve Bank, Philadelphia

Given the table we can compute a movement map.

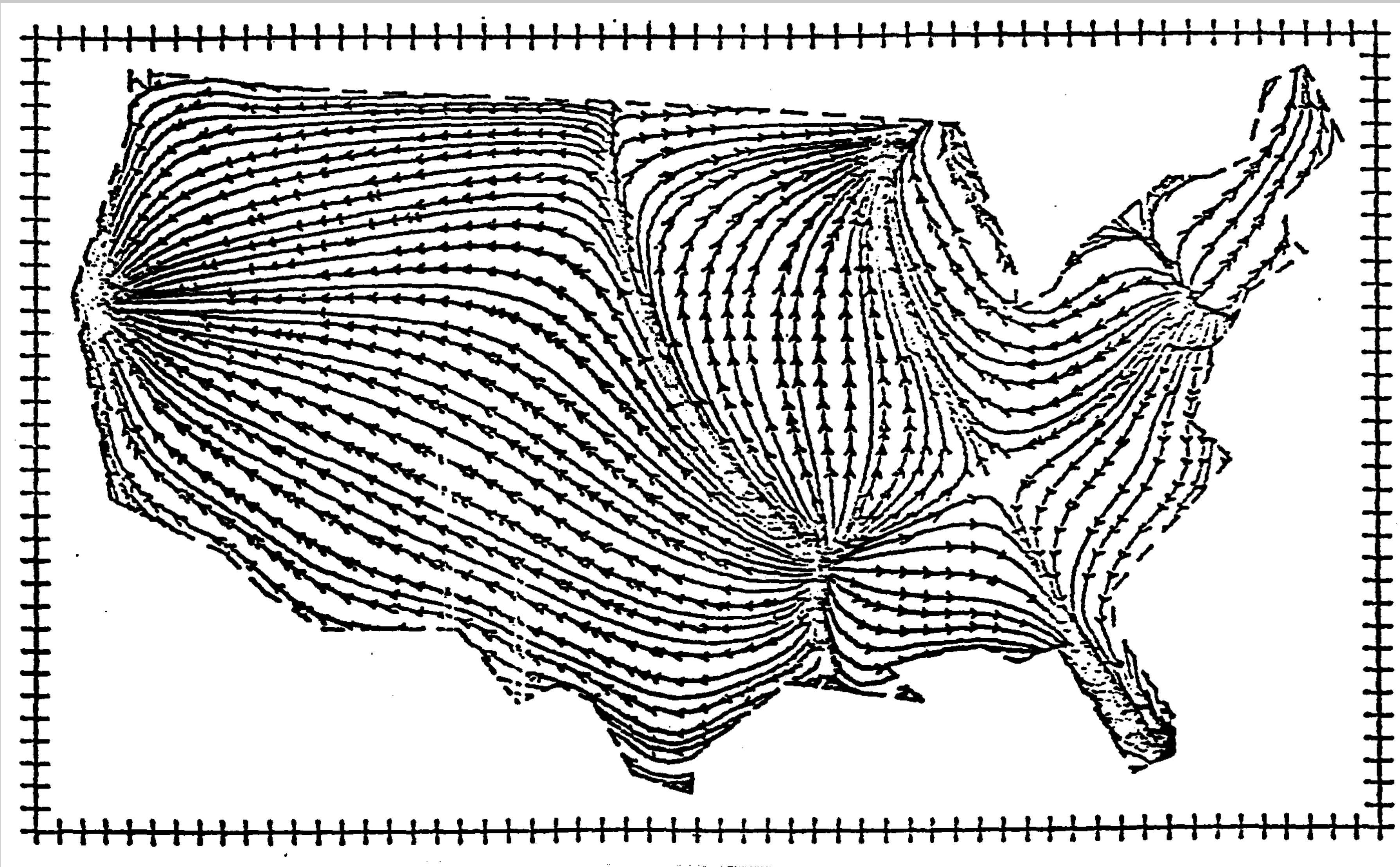
Movement of One Dollar Notes

between Federal Reserve Districts, in hundreds, Feb. 1976

From: \ To:	B	NY	P	Cl	R	A	Ch	SL	M	K	D	SF
Boston	2040	289	47	52	137	118	90	10	16	15	13	138
New York	602	1980	231	209	388	307	286	15	48	26	18	261
Philadelphia	143	414	860	84	342	130	134	8	25	10	10	80
Cleveland	68	192	47	1296	171	177	618	16	44	43	19	131
Richmond	150	266	158	226	3899	578	295	20	62	54	22	152
Atlanta	122	159	57	186	319	3741	439	30	51	78	102	189
Chicago	97	155	39	496	143	266	5630	74	278	100	40	290
St. Louis	31	56	14	142	80	201	573	342	46	128	47	109
Minneapolis	14	26	11	32	29	41	295	10	1438	51	14	138
Kansas City	20	41	8	55	40	71	215	33	120	811	86	247
Dallas	31	41	8	38	46	165	125	20	37	253	788	203
San Francisco	82	81	23	84	114	106	251	22	127	128	43	5380

Dollar Bill Movement in the U.S.

Showing the spatially coherent structures



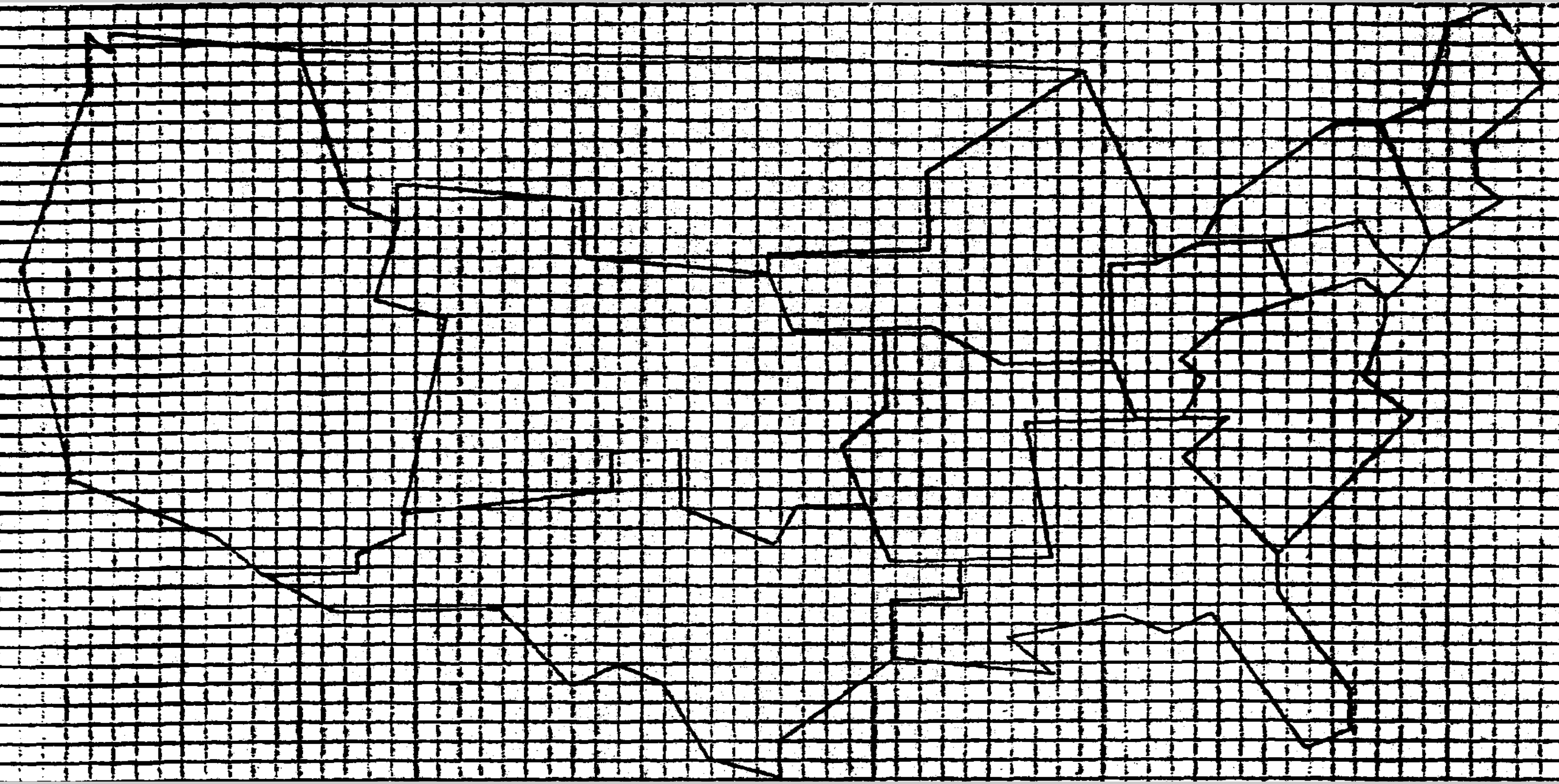
The map is computed using the continuous version of the gravity model

The result is the system of partial differential equations and solved by a finite difference iteration to obtain the potential field.

This can be contoured and its gradient computed and drawn on a map.

W. Tobler, 1981,"A Model of Geographic Movement", *Geogr. Analysis*, 13 (1): 1-20
G. Dorigo, & Tobler, W., 1983, "Push Pull Migration Laws", *Annals*, AAG, 73(1):1-17.

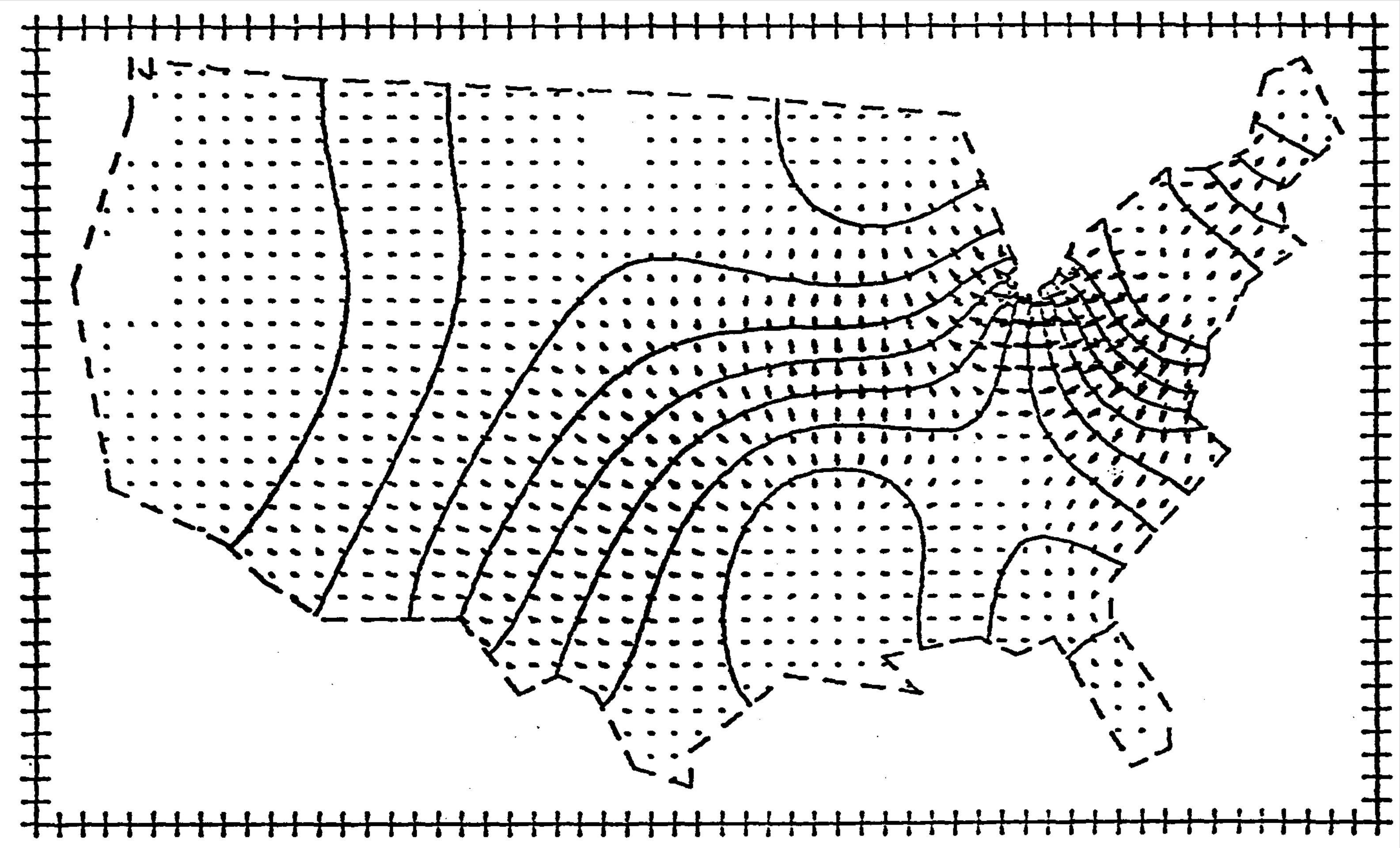
First the Federal Reserve Districts Are "Rasterized"



There will be one finite difference equation for each node on this raster

Solving the 2000 equations yields the potential

Shown here by contours and gradients.



The raster is indicated by the tick marks. The arrows are the gradients to the potentials. The earlier streakline map is obtained by connecting the gradient vectors.

Fifth Visualization

Biproportional Adjustment

Start with a square array, such as a migration table.

Such a table is asymmetric and the marginal values (row and column sums) are unequal.

They can be rendered equal by an iteration (IPFP).

This has the effect of removing size differences.

I have always wondered what the IPFP does to the data. Haven't you?

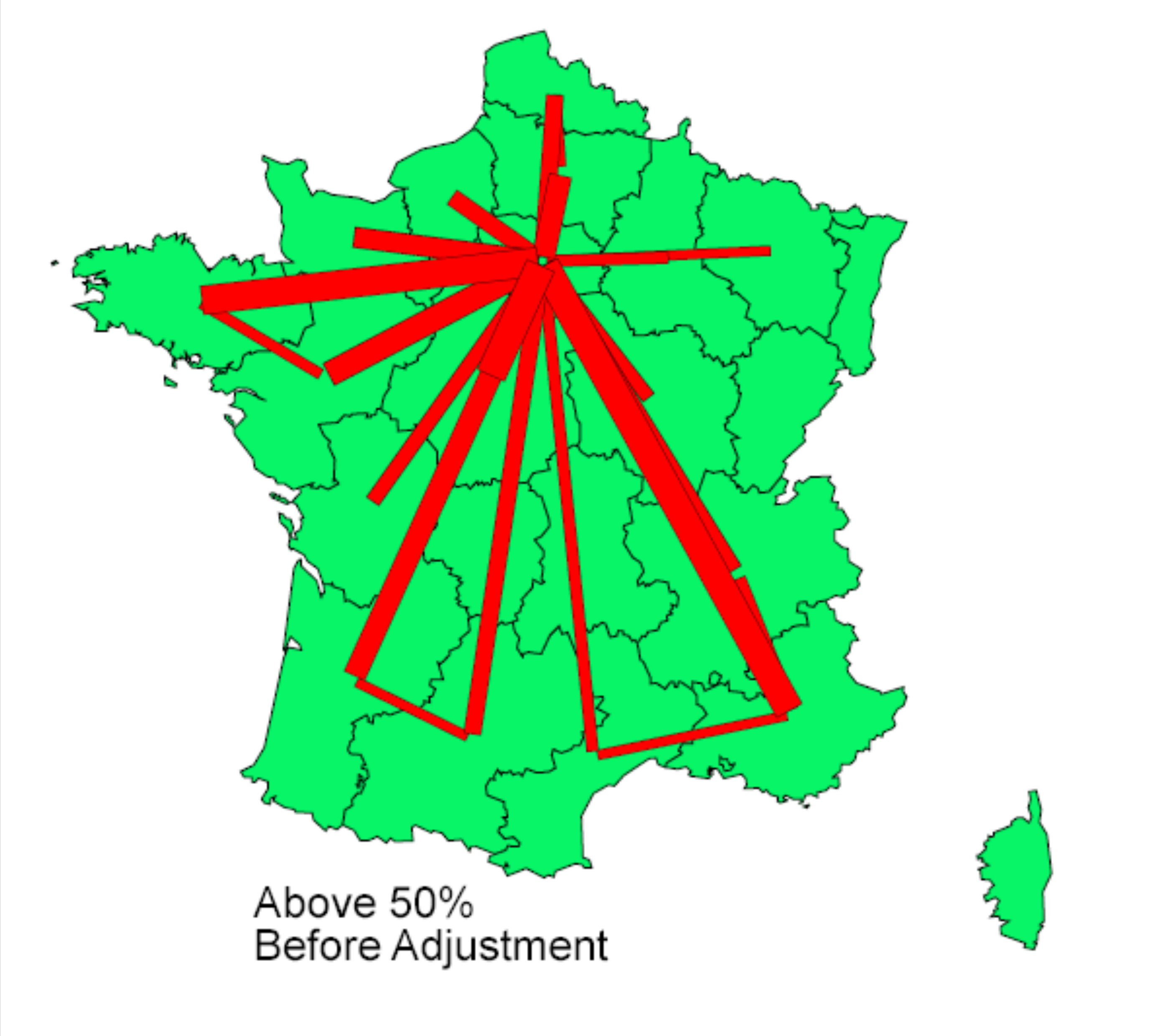
My first example is based on migration of employed Frenchmen 1962-1968 (INSEE).

Data source:

Rouget, 1972, "Graph Theory and Hierarchisation Models" *Regional and Urban Economics*, 2,3 :263-296

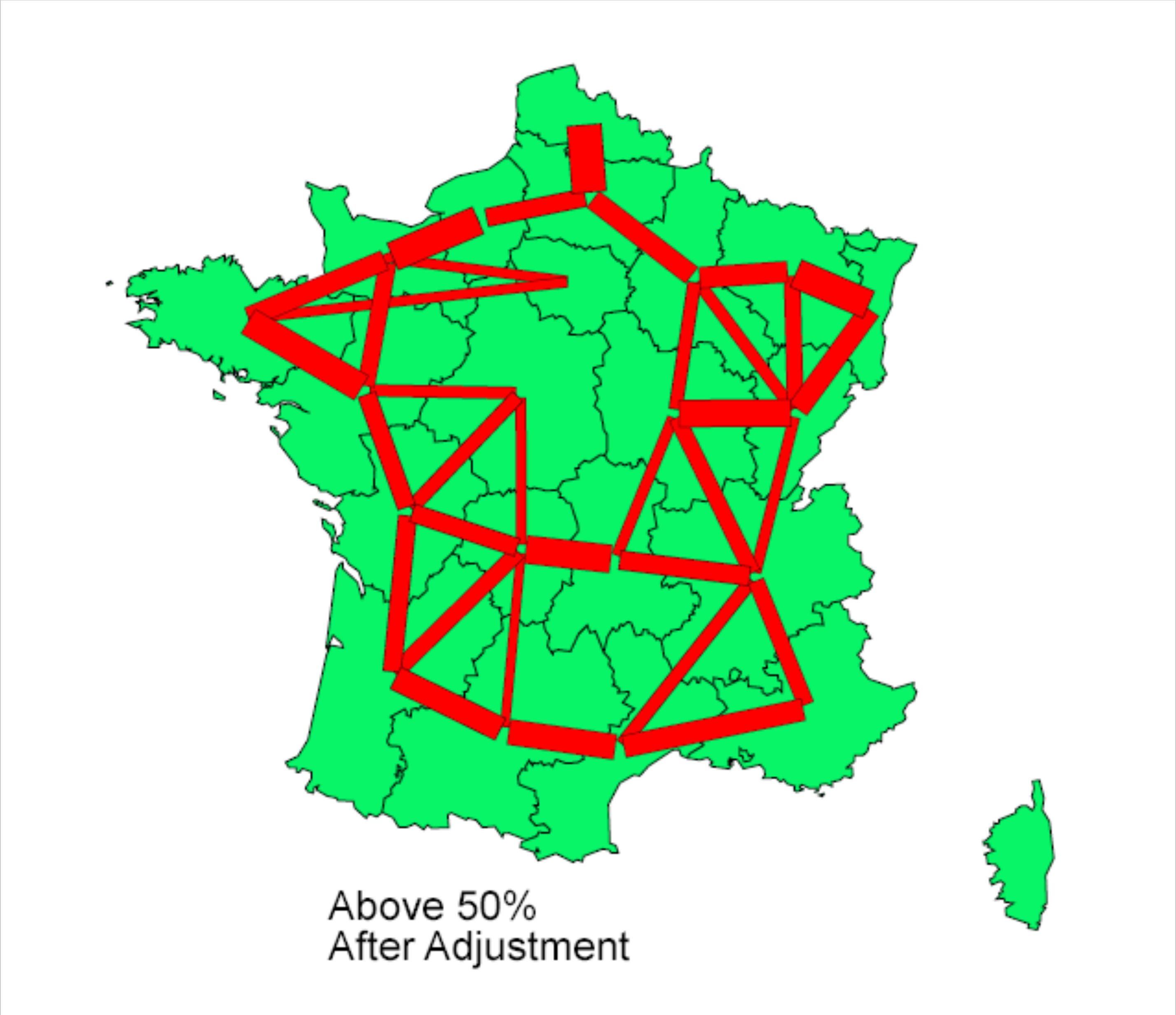
Slater, 1976, "Hierarchical Internal Migration Regions of France", *IEEE Transactions on Systems, Man, and Cyrbernetics*, 321-324.1

Paris dominates before the adjustment



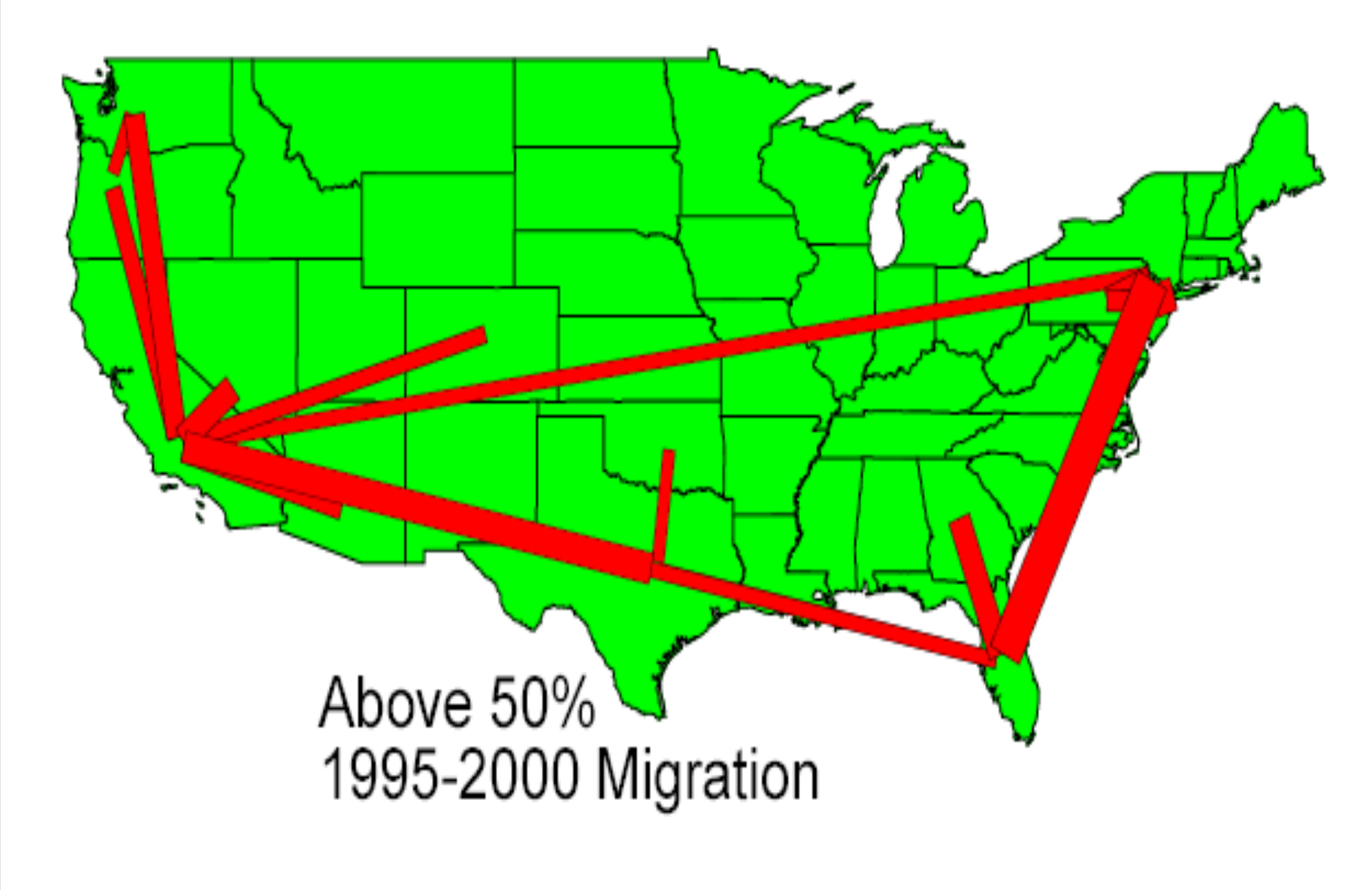
What biproportionalizing has done to the data

Adjustment to equal marginal values.



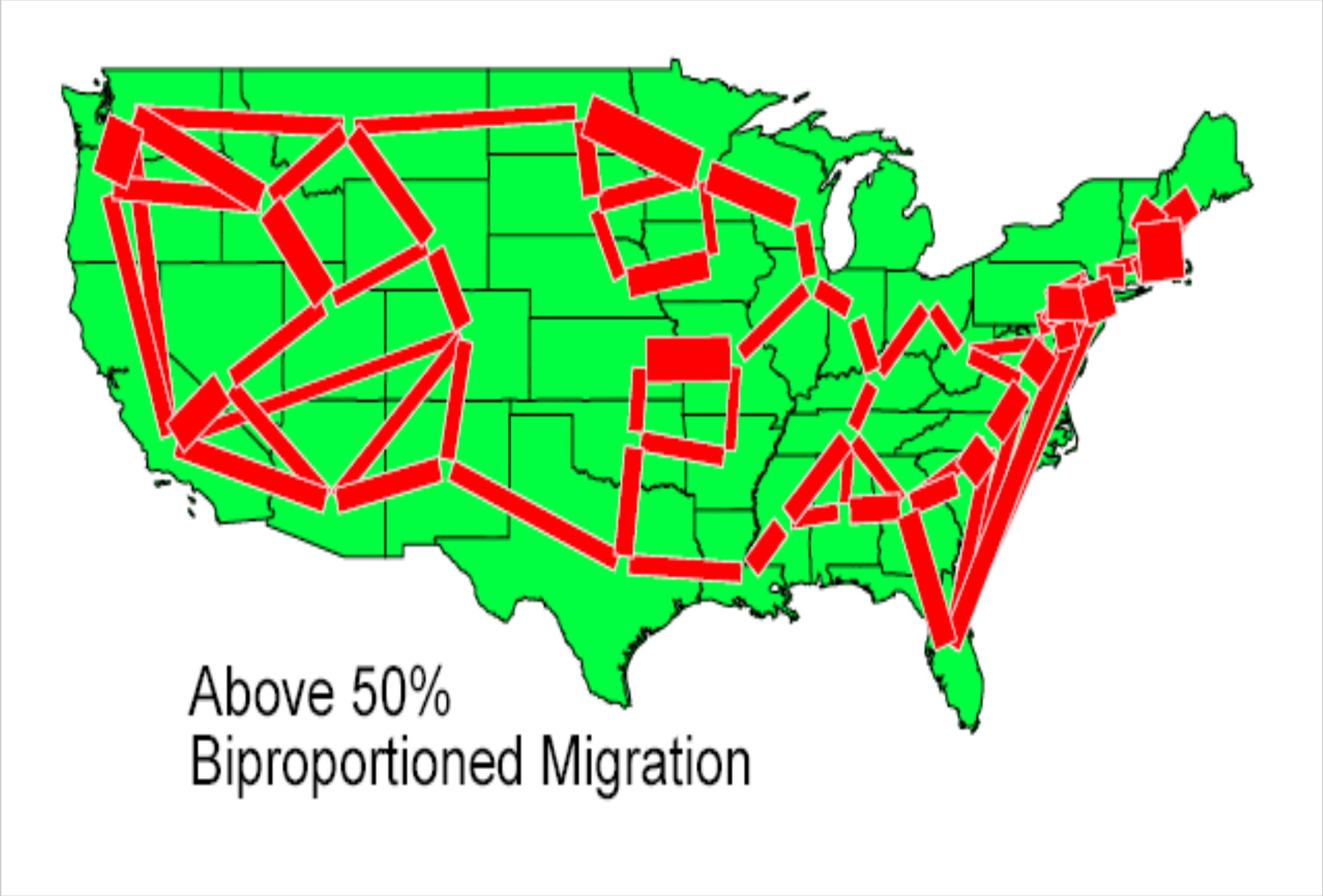
Another Example

US Migration before adjustment.



When the array is modified by the Iterative Proportional Fitting Procedure the smaller values also appear.

Adjustment to equal marginal values.



Sixth Visualization

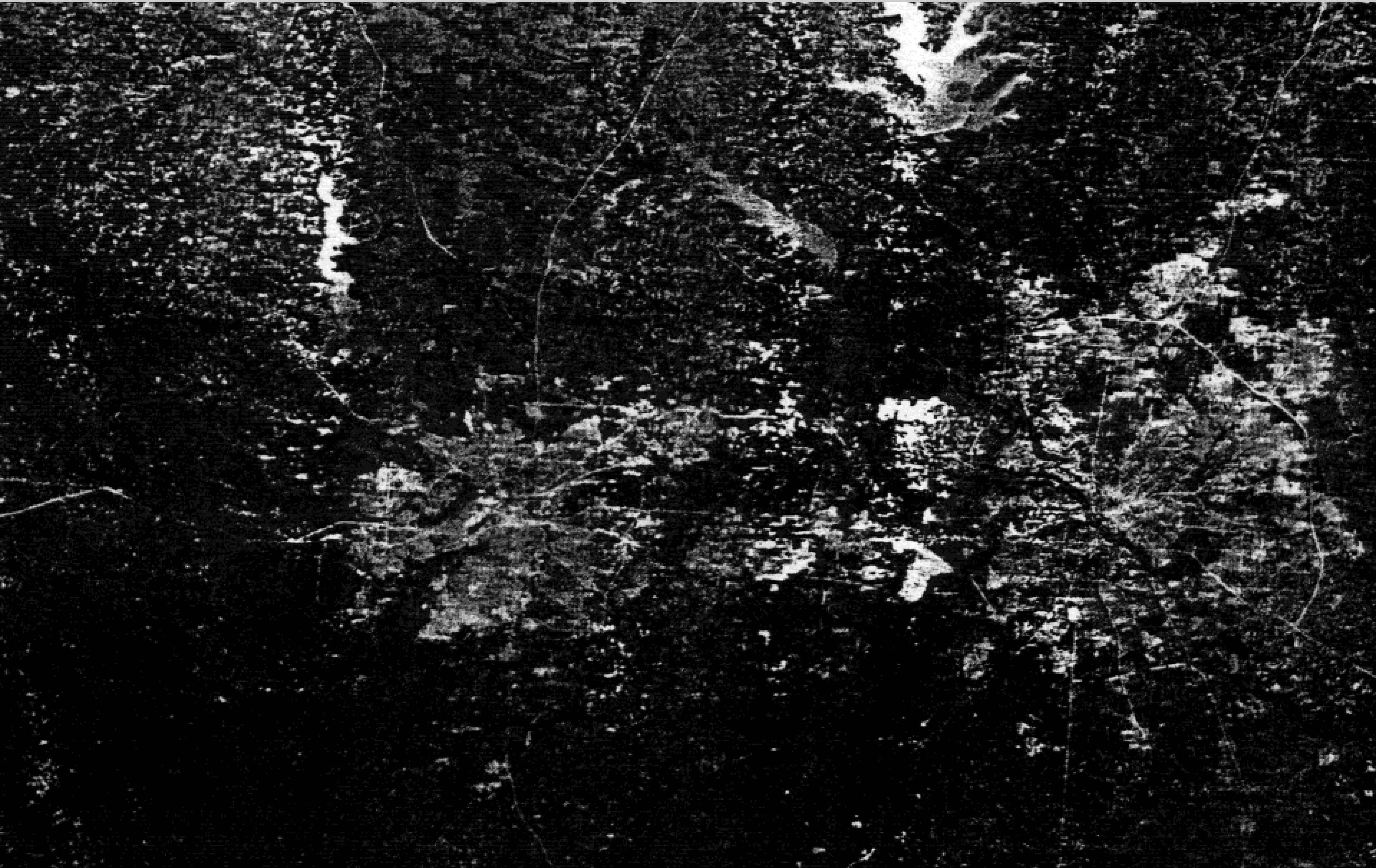
Coalescent Cities

A satellite image of adjacent cities,
providing visual evidence for urban potential fields,
illustrates the dynamics of urban growth.

Appolo VI Hasselblat photograph AS6-2-1462.
W.Tobler, “Visual Evidence for Urban Potential Fields”, *Mappemond*, 1991,1: 46-47

Forth Worth and Dallas

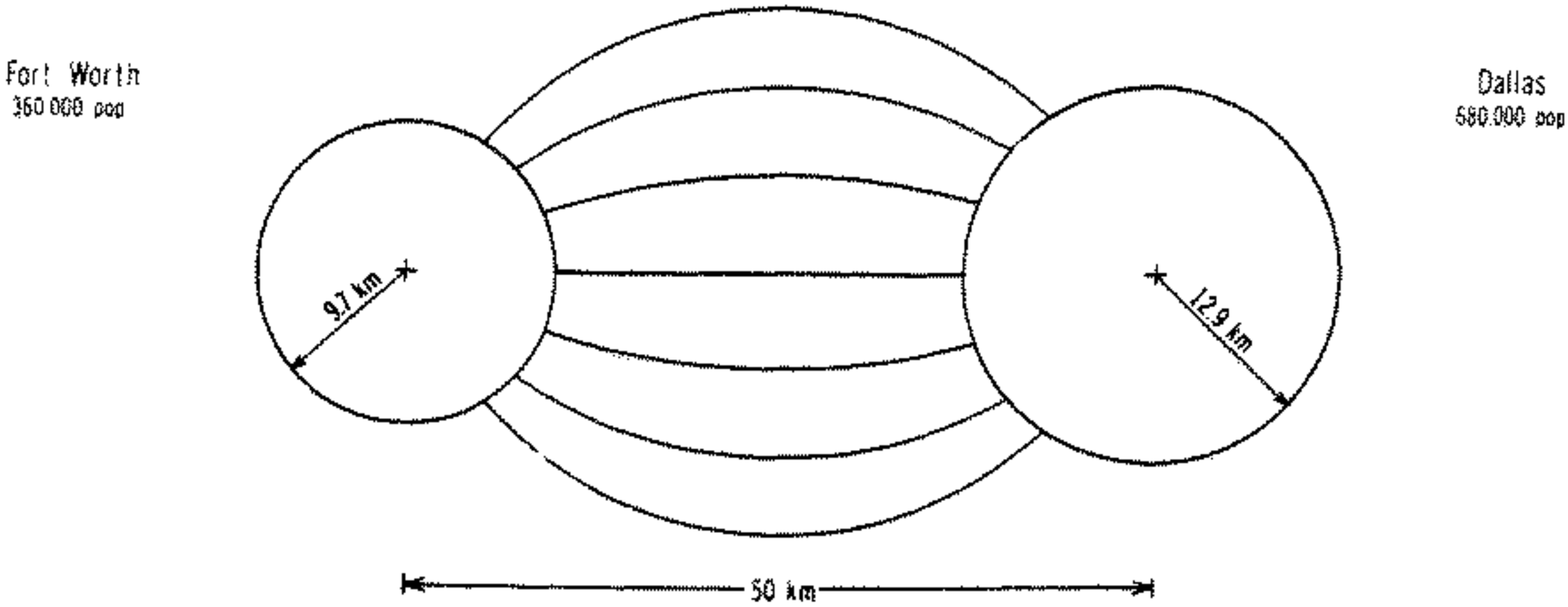
Growing towards each other.



For comparison a schematic diagram with a portion of the expected stream function

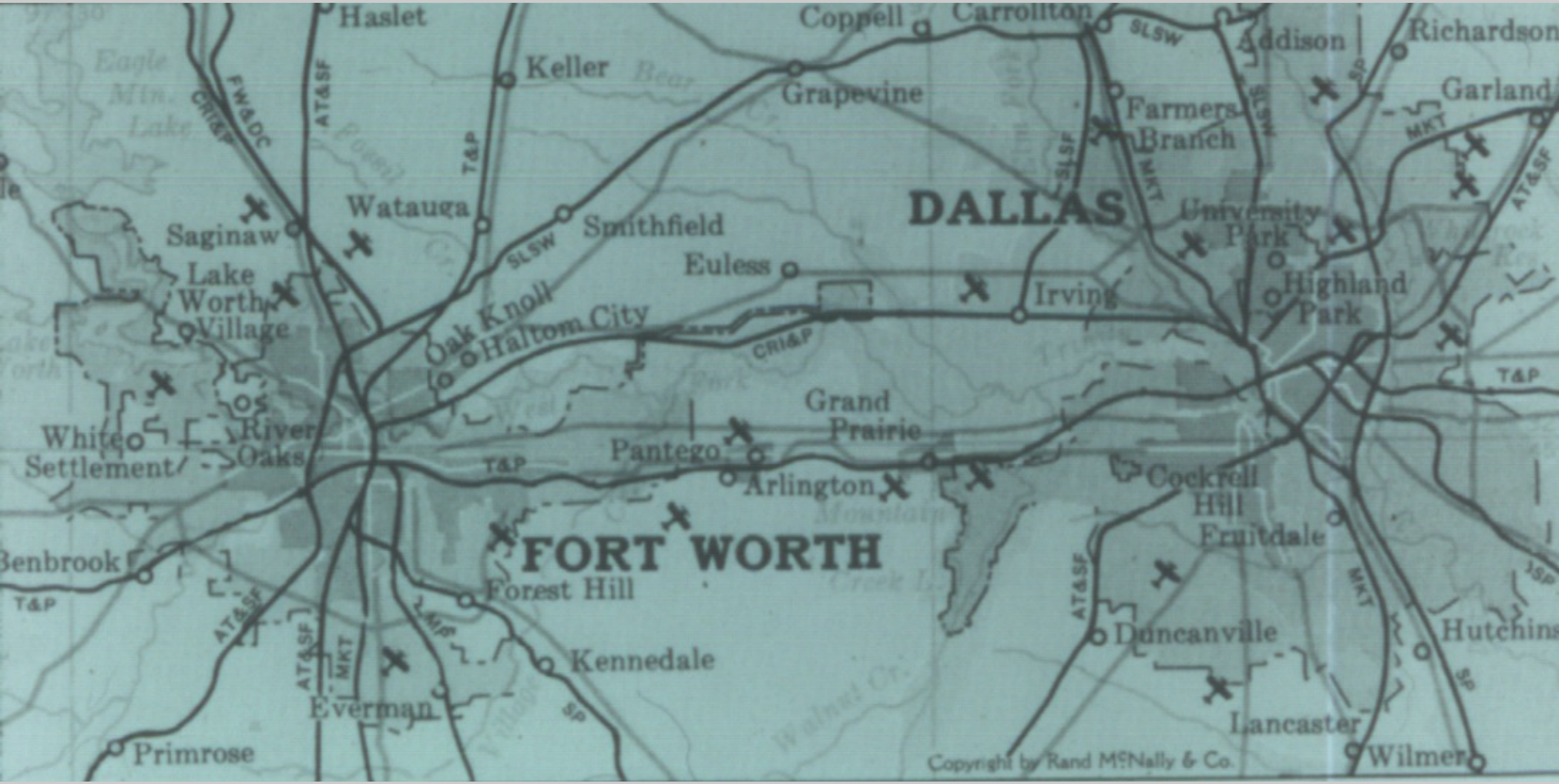
if Fort Worth and Dallas induce a potential field.

Circle area based on population. See: W. Tobler, "Satellite confirmation of settlement size coefficients", *Area.,* 1969, 1,3:30-34.



A more recent image of the area.

And a map of the area.



Looking for more examples

Potential examples - some California speculations

San Luis Obispo - Santa Maria

Ventura - Oxnard - Simi Valley

Sacramento - Stockton - Modesto

Los Angeles - Orange County - San Diego

Los Angeles to San Diego

More possibilities?

Seattle-Tacoma

Millwaukee - Chicago

Washington D.C. - Baltimore

Examples in Asia?

South America, Europe, Africa?

Topics for student papers?

Seventh Visualization

Multidimensional scaling

(a.k.a. MDS)

The problem is to find relative locations when given dissimilarities taken to be distances.

That is, given D_{ij} find X and Y when $D_{ij} = [(X_i - X_j)^2 + (Y_i - Y_j)^2]^{1/2}$

In high school one learns to compute distances from coordinates.

Why does one not learn the reverse, compute coordinates from distances?

One mathematical procedure is as follows:

1: Guess

Much of the Angst regarding mathematics is because guessing is not taught as a legitimate mathematical procedure.

2: Systematically improve the guess by taking into account the given information.

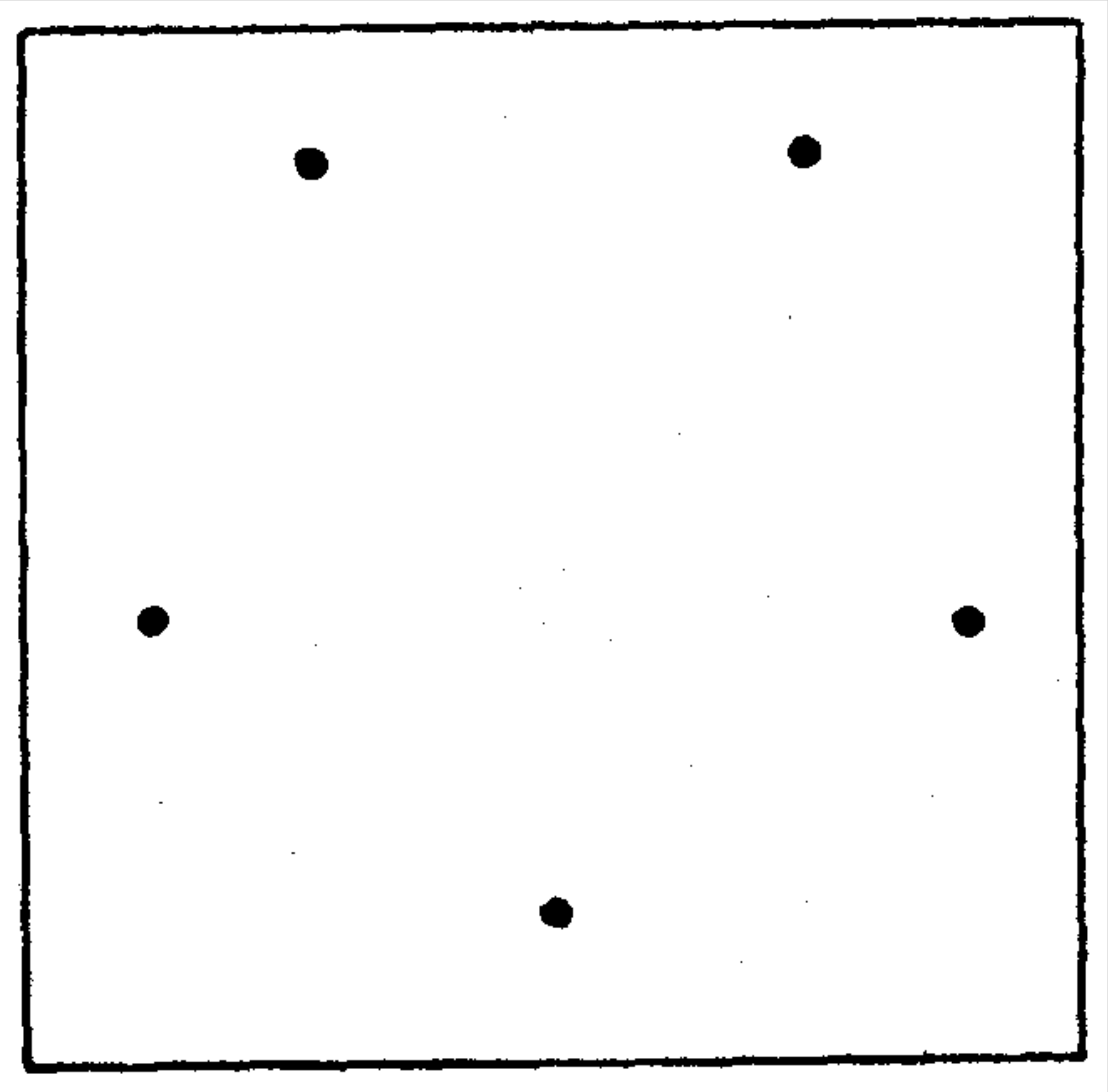
This can be visualized in the following example.

A Classic Example

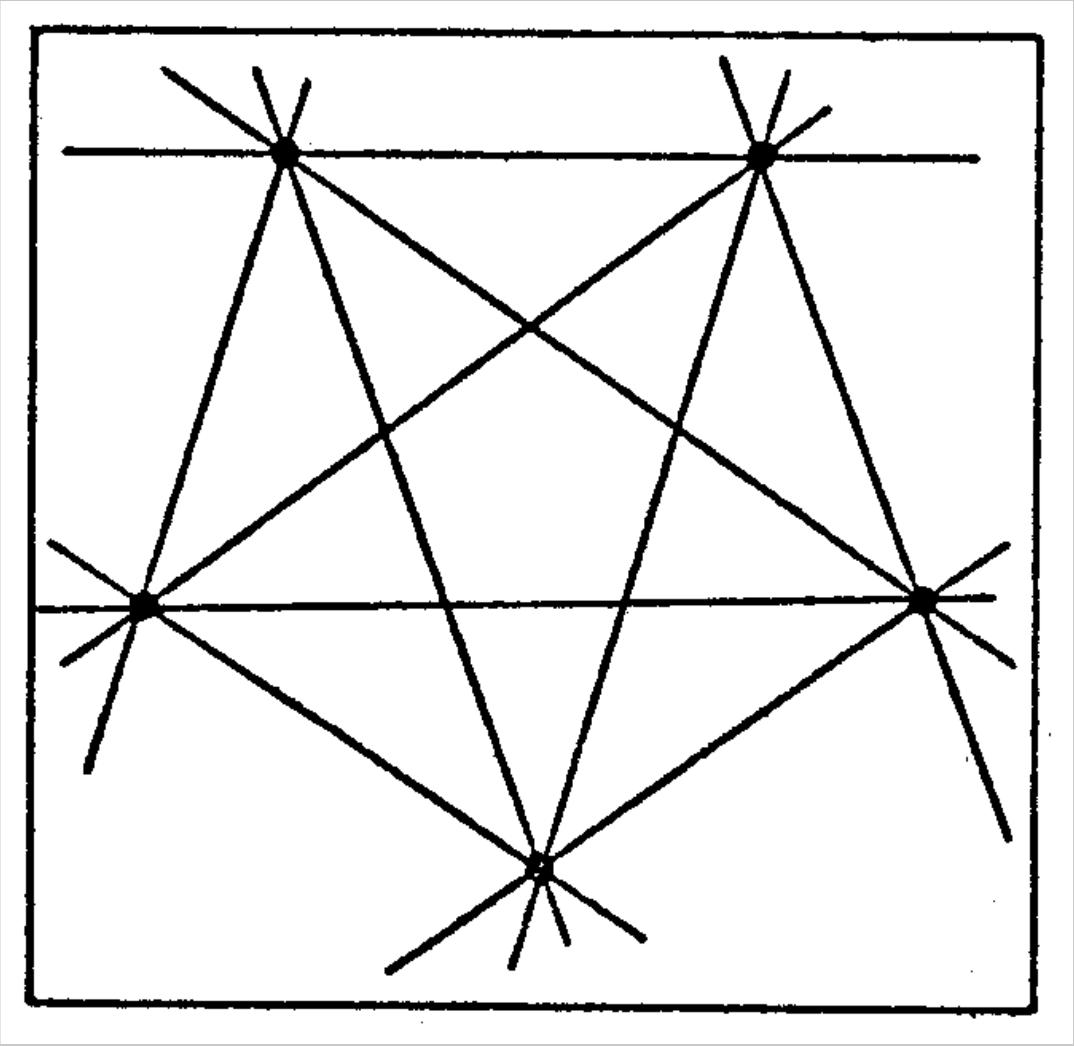
Use the road distances to make a map showing the locations.

ATL	0								
BOS	1100	0							
DAL	810	-	0						
DEN	1440	1990	800	0					
LAX	-	-	1430	1170	0				
MIA	680	-	1330	2090	-	0			
MIN	1130	1400	-	870	-	-	0		
SFO	-	-	1770	1270	400	-	2000	0	
SEA	-	-	-	1370	-	-	1640	820	0
	ATL	**BOS**	**DAL**	**DEN**	**LAX**	**MIA**	**MIN**	**SFO**	**SEA**

Step 1: Guess at locations.

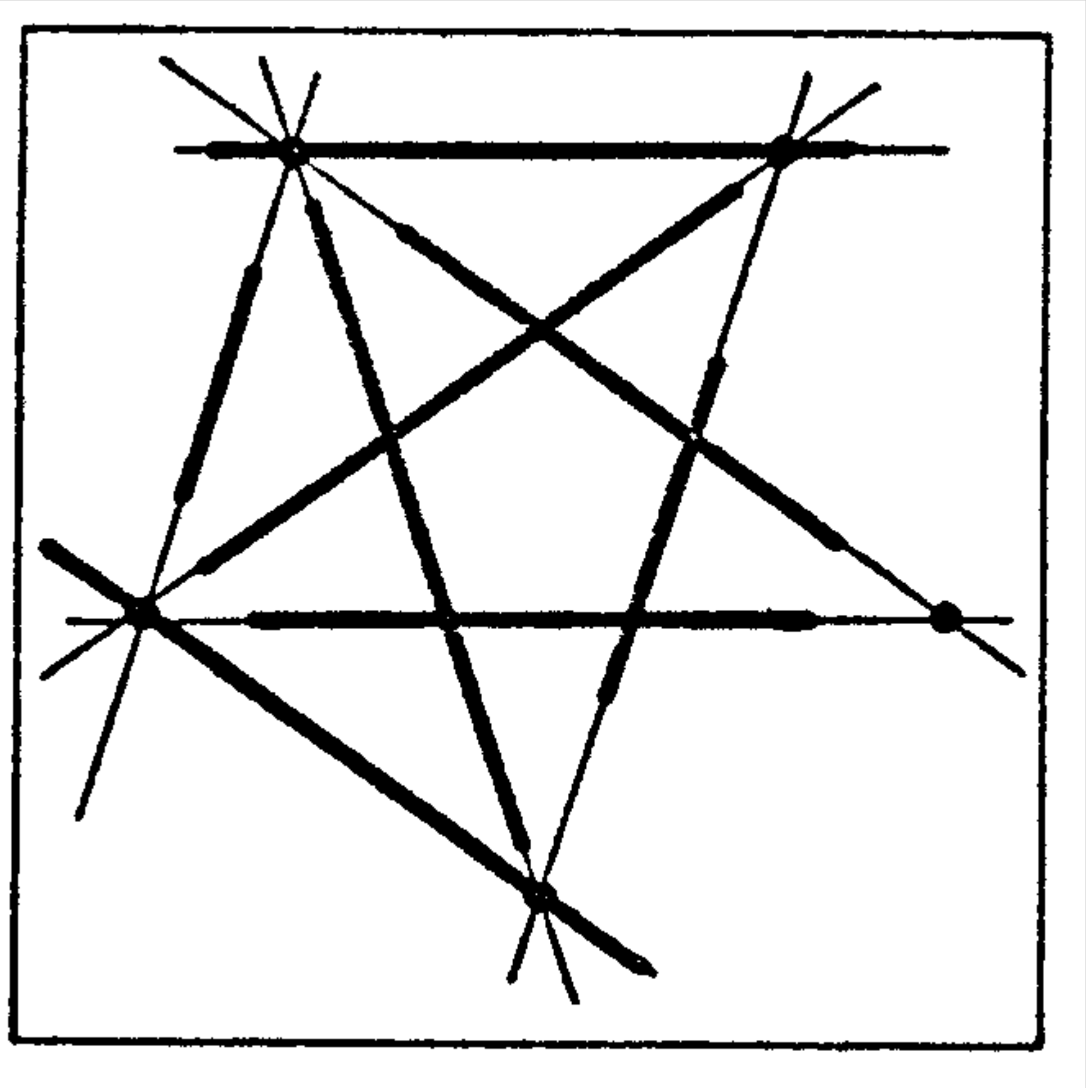


Step 2: Insert lines to represent distances.

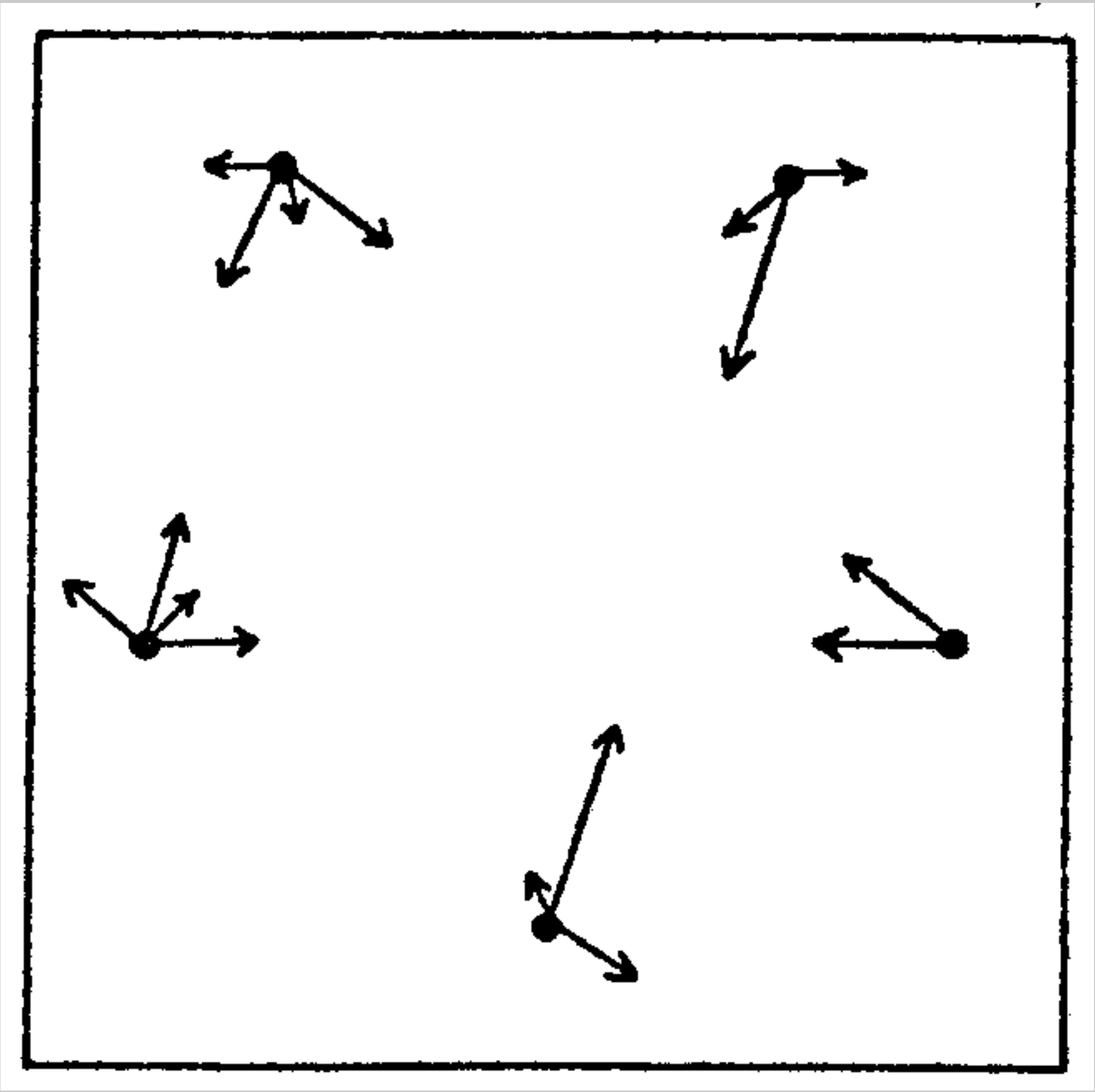


Step 3: Insert estimated distances.

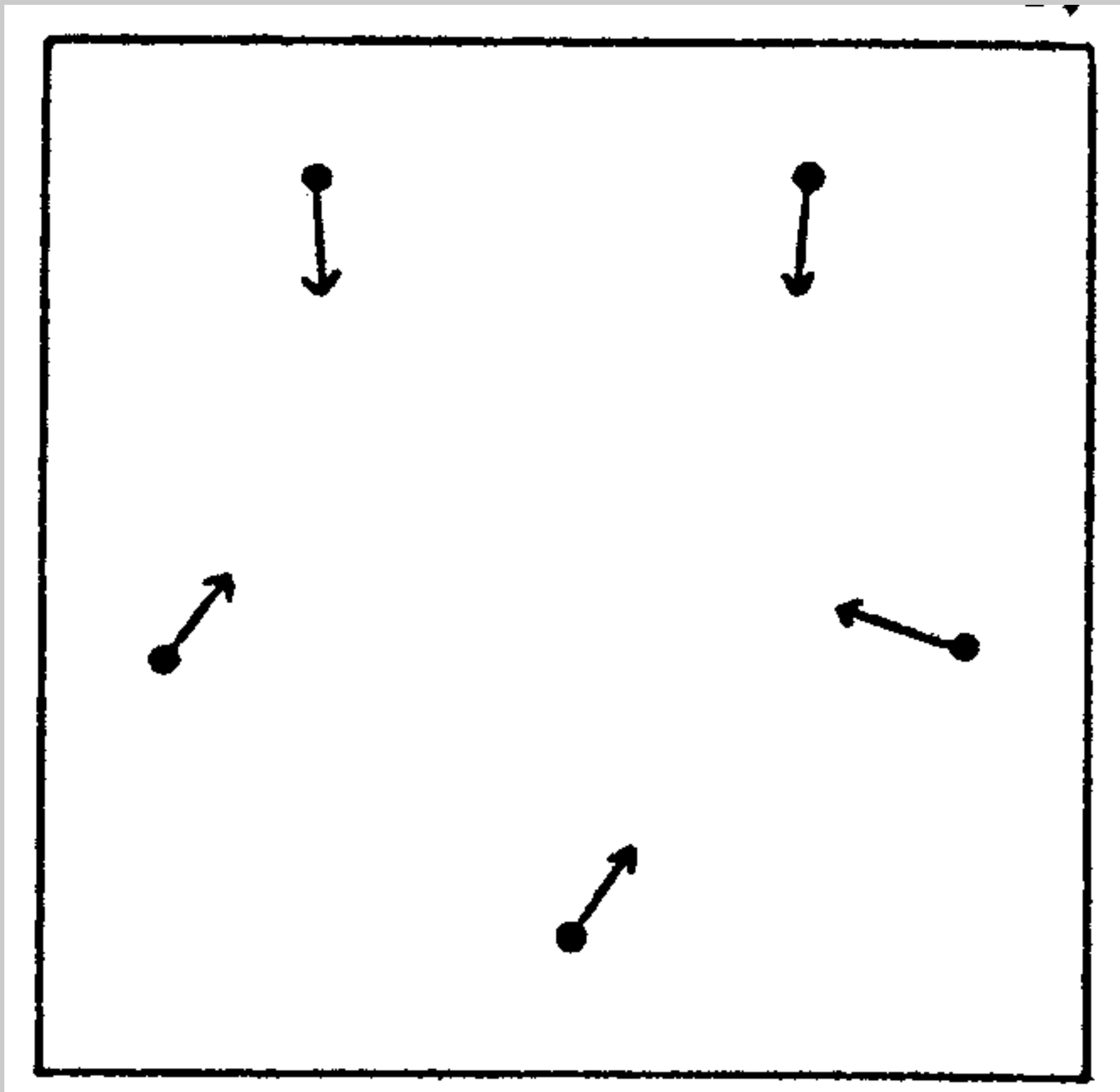
Some will be too short, some too long, some missing.



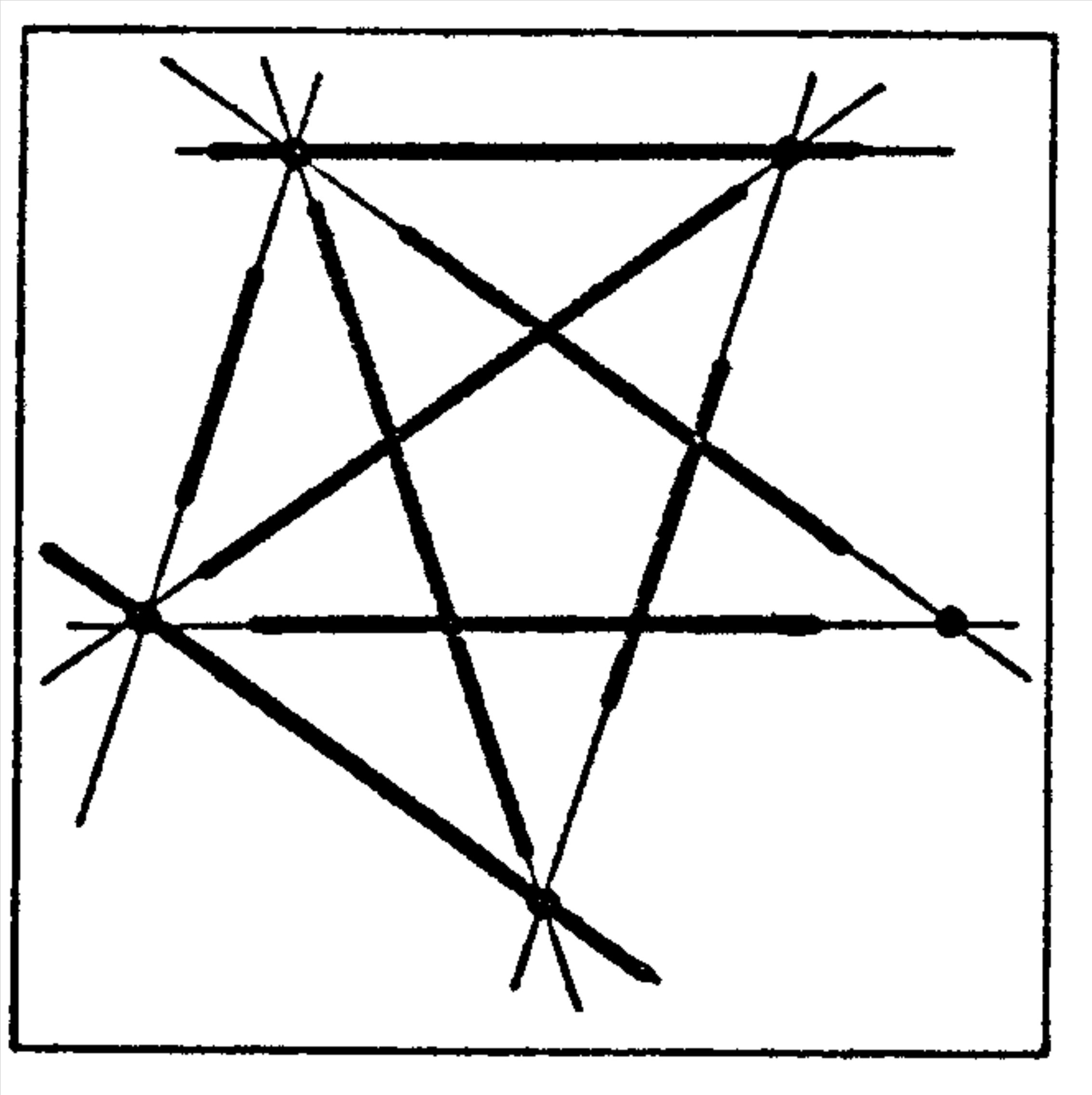
Step 4: Examine and average discrepancies.



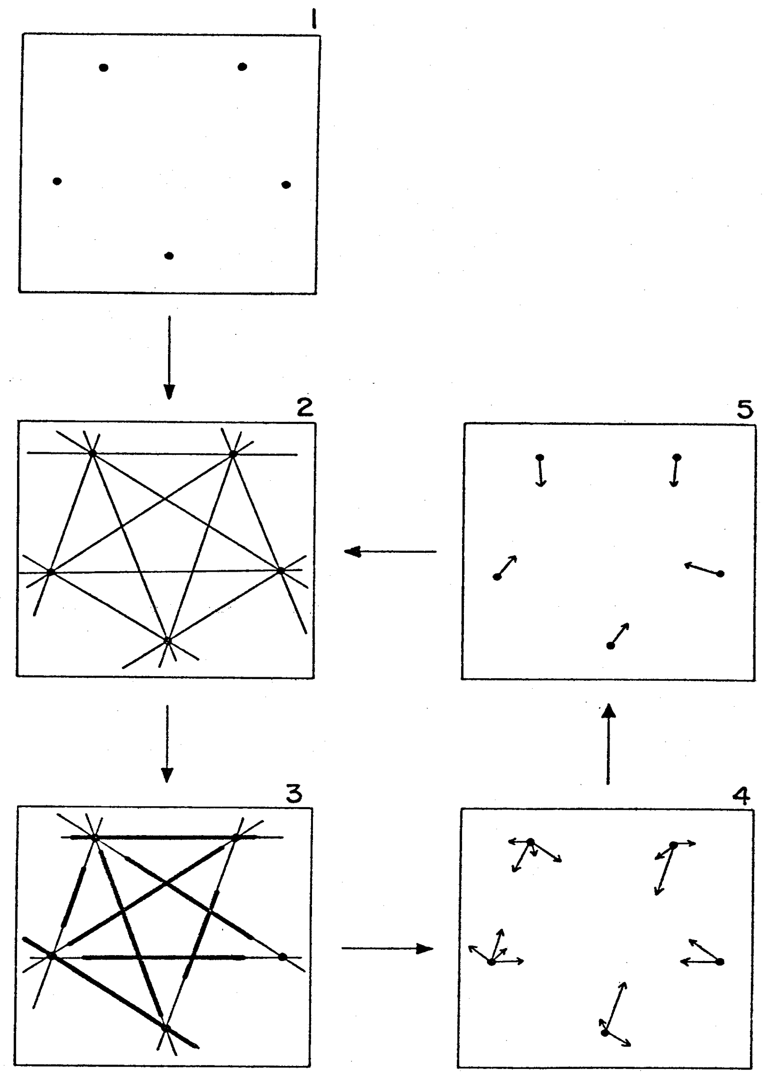
Step 5: Apply adjustments.



Step 6: Insert adjusted distances and repeat.



The complete sequence of iterative steps.



For the USA data

The first step is to systematically array the data in a table; only half of the table is needed, since it's symmetric, and places with only one connection can be ignored:

ATL	0								
BOS	1100	0							
DAL	810	.	0						
DEN	1440	1990	800	0					
LAX	.	.	1430	1170	0				
MIA	680	.	1330	2090	.	0			
MIN	1130	1400	.	870	.	.	0		
SFO	.	.	1770	1270	400	.	2000	0	
SEA	.	.	.	1370	.	.	1640	820	0
	ATL	BOS	DAL	DEN	LAX	MIA	MIN	SFO	SEA

Only twenty of the possible thirty six distances are given. As initial coordinates one can use:

	X	Y
1	2600	700
2	3300	1800
3	1800	500
4	1050	1050
5	300	500
6	3100	150
7	1900	1700
8	100	900
9	200	1700

obtained by placing the cities in their approximate location on millimeter paper.

The answer, with estimated coordinates and distances.

After several hundred iterations the root mean square error is 44.4 units (a 99.1% fit to the data) and the solution coordinates are:

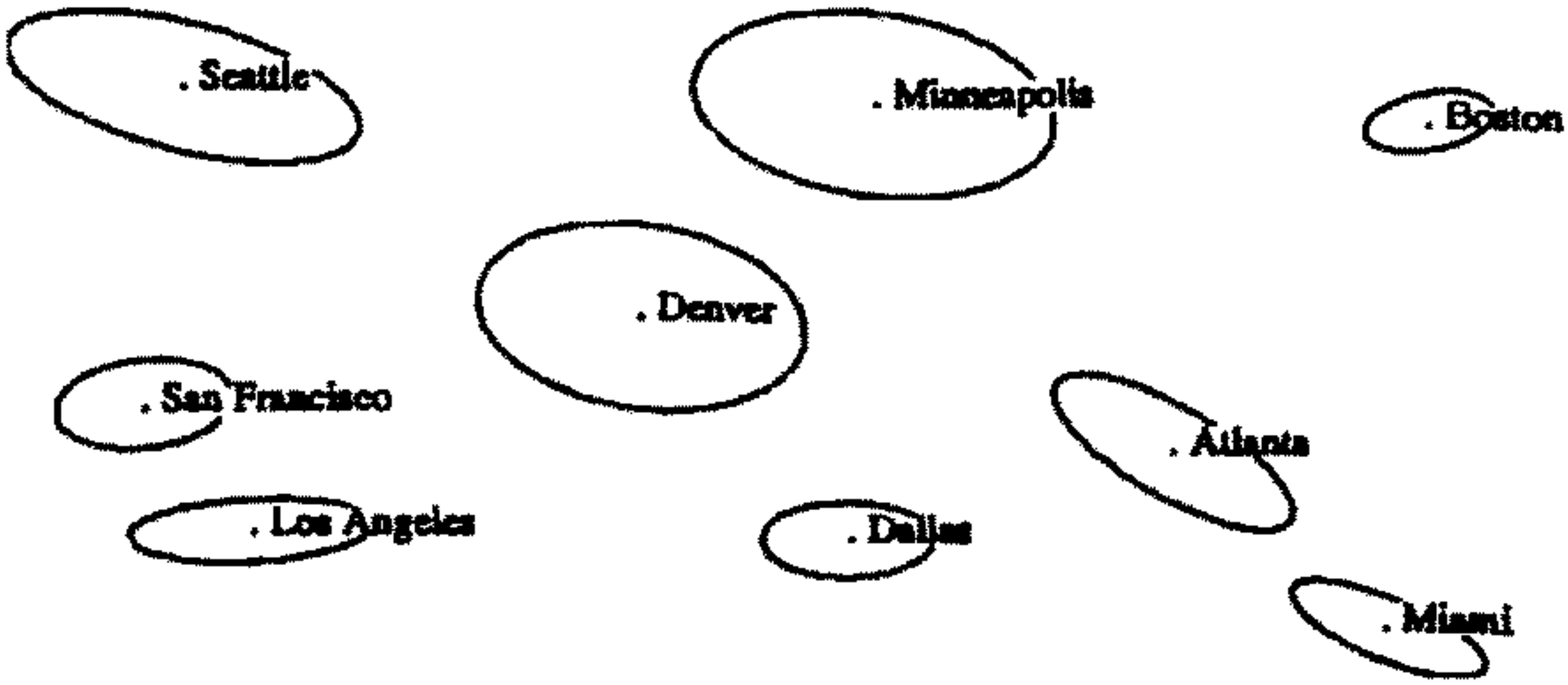
	X	Y
1	2538	700
2	3168	1588
3	1727	473
4	1211	1073
5	242	497
6	3066	233
7	1794	1637
8	-19	837
9	79	1689

Resulting distances are, to the nearest mile,

	ATL	BOS	DAL	DEN	LAX	MIA	MIN	SFO	SEA
ATL	0								
BOS	1089	0							
DAL	842	1822	0						
DEN	1378	2023	791	0					
LAX	2304	3122	1485	1127	0				
MIA	705	1359	1361	2036	2836	0			
MIN	1196	1375	1166	811	1925	1895	0		
SFO	2560	3274	1783	1252	429	3143	1981	0	
SEA	2650	3090	2048	1289	1202	3323	1715	857	0

The solution

Showing (exaggerated) standard errors.



USA Highway Distance Map

Here is a more elaborate example.

Using a road atlas the student took many values from the table of distances between places. These tables are common in such atlases.

Using these distances he then computed the location of the places.

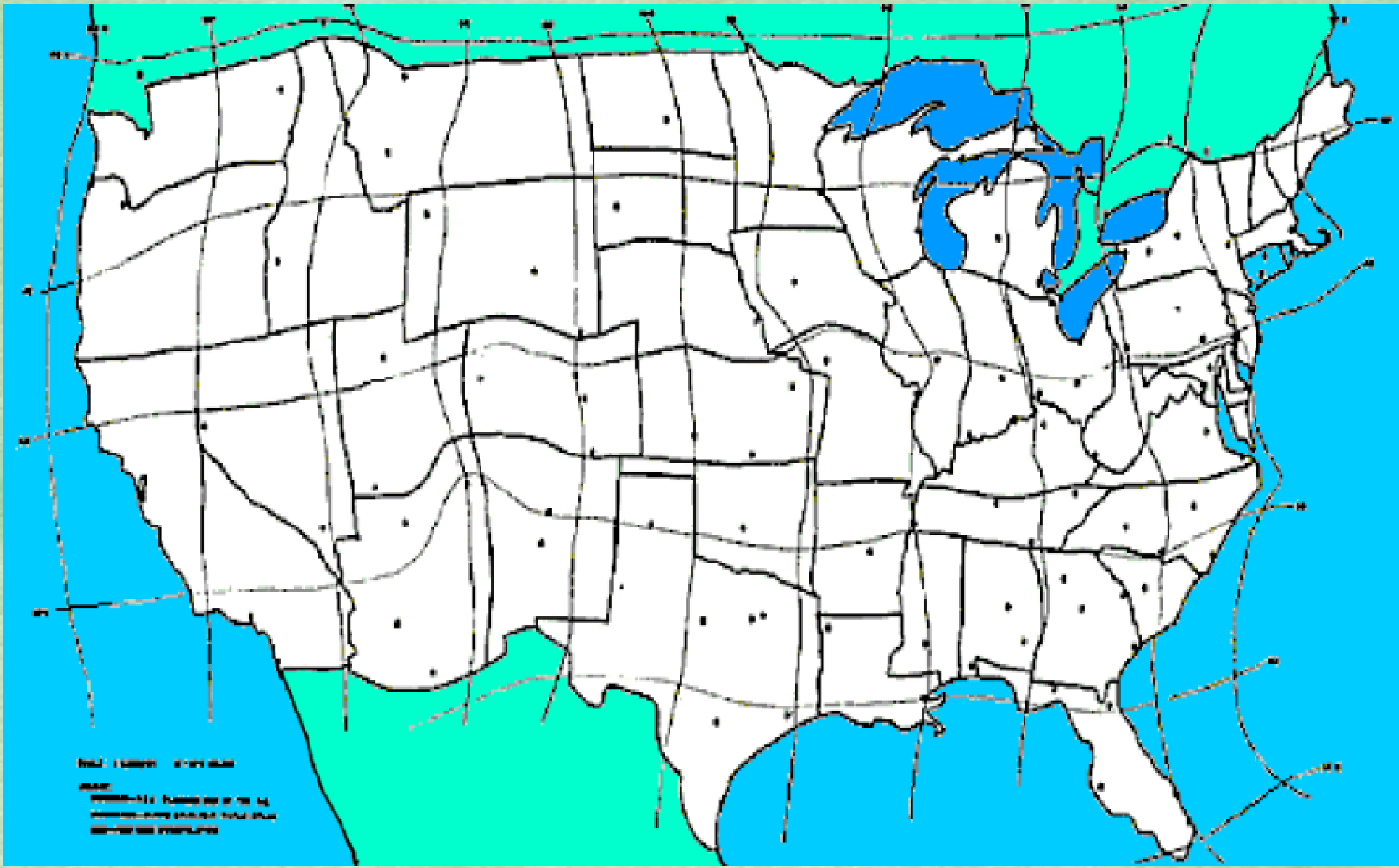
The US outline and latitude longitude grid were then interpolated to complete the map.

As for all maps, Tissot's indicatrix can be used to measure the linear, angular, and areal distortion

in this case introduced by the roads.

Road Distance Map of the United States

Student drawing



Another example

Cappadocian Speculation

Given hundreds of cuneiform tablets from 1700 B.C.
Use the ones that mention places in Cappadocia (Central Turkey).

To find the location of the places.

Tobler & Weinberg, *Nature,* 1971 (May 7), 231: 39-41

Cappadocian Cuneiform Tablet

Some giving trade between places

About 5 by 8 cm

Some tablets include the names of towns and occasionally indications of trade.

The frequency count of town names can be taken to be proportional to their size or population.

When two town names occur on one tablet this implies an interaction between them.

So we have P_i and P_j and $T_{ij.}$
in this case relating 62 towns.

Use the gravity model

T = trade, P = population, D = distance

$$T_{ij} = k\, P_i\, P_j \,/\, D_{ij}$$

An exponent on the distance can also be used, with obvious modification.

Other model variants could also be tried.

Invert the model to get

$$D_{ij} = k\, P_i\, P_j \,/\, T_{ij}$$

Given distances compute the locations.

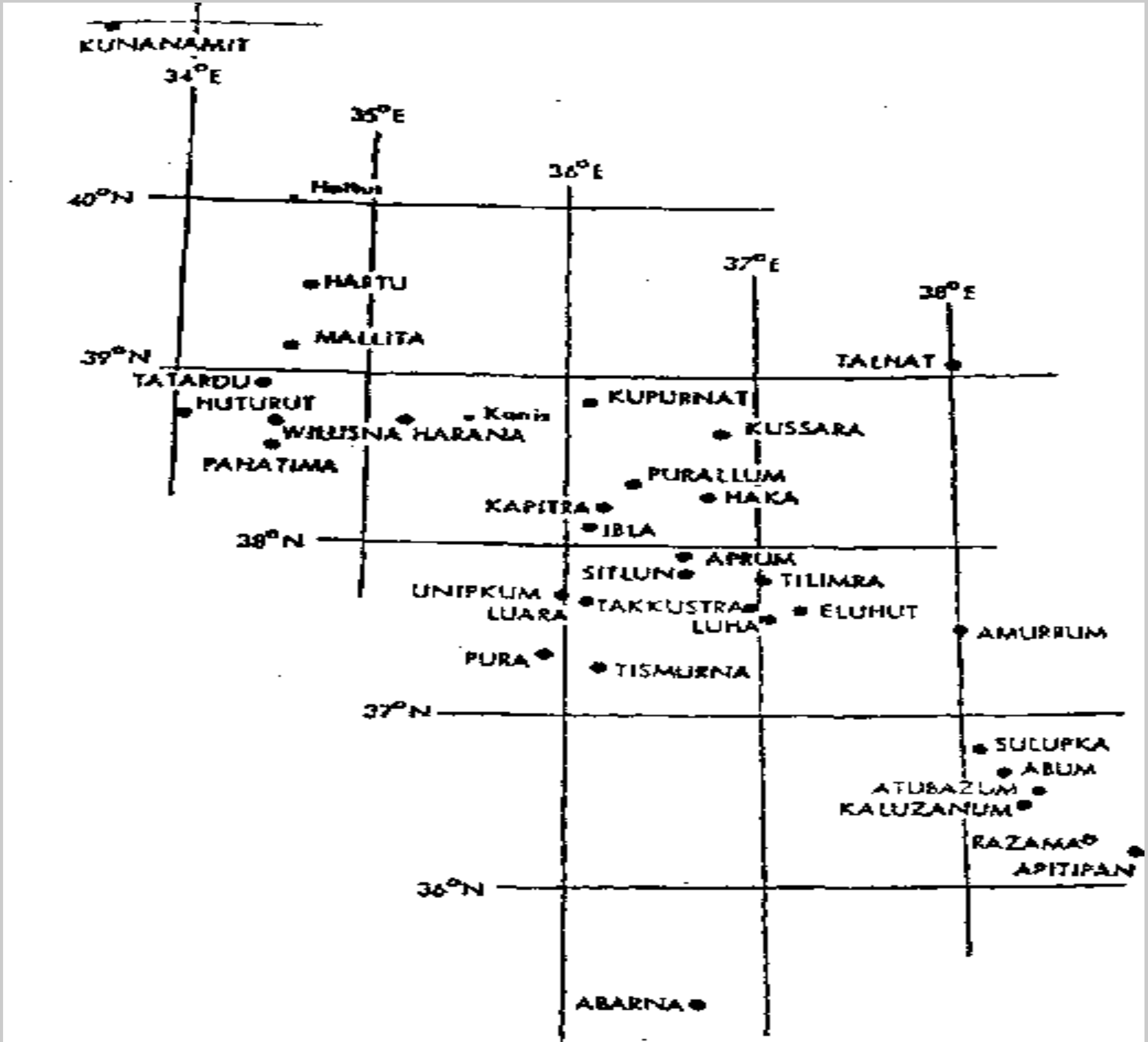
That is, find the latitudes and longitudes

using the (rather sparce) symmetric 62 by 62 table of

distances and the very few known locations.

Predicted locations

For standard errors see the published text



Another example

Using an input/output table, consider two industries to be “close” if the quantity of exchanges between them is large.

Construct a diagram in which “close” industries are placed near each other, as follows.

Place the 1st industry at random, Then the next “closest” on a circle of chosen radius, then the 3rd where circles about the first two intersect. The 4th industry goes where more circles intersect.

Obviously this gets very tedious and can’t go on.

So use the iterative procedure to place “close” industries near each other.

Part of a 50 by 50 Input - Output Table

How can one comprehend such a large table without long and tedious study?

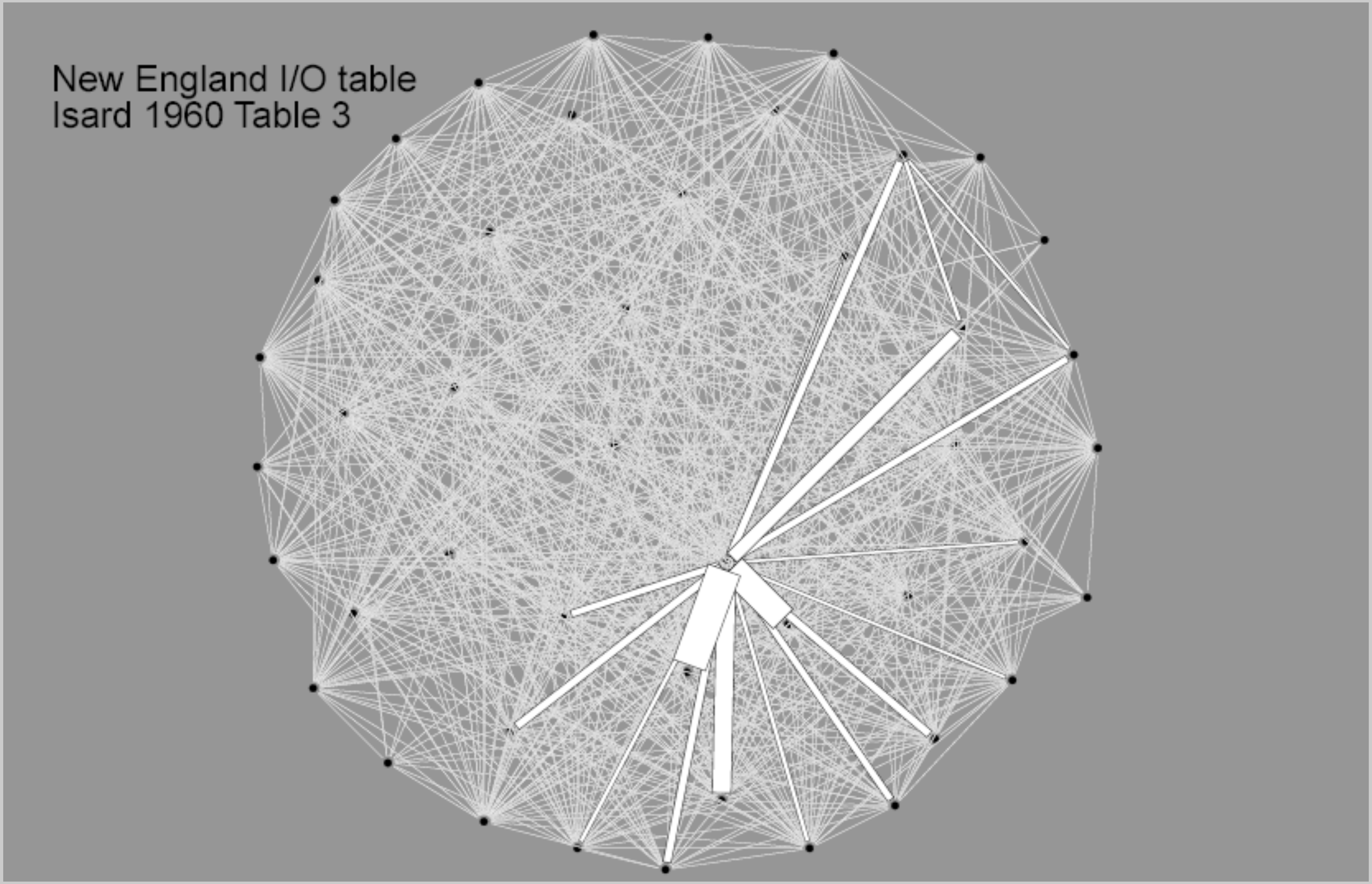
Industry (Sector) Producing \ Industry (Sector) Purchasing	1. Agriculture and Fisheries	2. Food and Kindred Products	3. Tobacco Manufactures	4. Textile Mill Products	5. Apparel	6. Lumber and Wood Products	7. Furniture and Fixtures	8. Paper and Allied Products	9. Printing and Publishing	10. Chemicals
1. Agriculture and Fisheries	244.5	534.2	7.6	571.2	1.2	8.3	–	1.2	–	57.9
2. Food and Kindred Products	53.6	174.3	0.1	16.5	0.6	*	*	4.1	*	32.8
3. Tobacco Manufactures	–	–	8.1	–	–	–	–	–	–	†
4. Textile Mill Products	1.4	0.1	–	358.0	246.0	0.1	15.4	5.9	1.6	0.6
5. Apparel	1.0	7.2	–	–	124.2	–	0.3	2.8	–	1.4
6. Lumber and Wood Products	3.3	2.9	0.2	4.9	0.1	47.2	20.8	36.8	0.1	2.2
7. Furniture and Fixtures	–	–	–	3.3	–	–	0.4	0.7	–	–
8. Paper and Allied Products	†	16.1	0.6	21.4	1.6	0.2	0.8	358.0	70.3	15.8
9. Printing and Publishing	–	1.4	–	0.6	–	–	–	–	49.8	0.8
10. Chemicals	18.7	51.5	0.2	219.8	9.0	1.1	3.4	25.2	6.3	127.0
11. Products of Petroleum and Coal	10.3	2.1	*	8.2	0.3	3.2	0.1	8.7	0.2	15.5
12. Rubber Products	2.7	0.3	–	3.6	1.1	0.4	0.3	1.2	0.2	†
13. Leather and Leather Products	–	–	–	0.6	3.4	0.2	0.4	–	0.3	–
14. Stone, Clay and Glass Products	1.5	9.0	†	0.3	–	0.6	1.8	3.9	–	12.3
15. Iron and Steel	0.1	0.1	–	–	0.1	0.4	5.2	–	–	0.2
16. Nonferrous Metals	–	–	–	–	–	0.1	0.9	–	0.9	9.0
17. Plumbing and Heating Supplies	–	–	–	–	–	–	–	–	–	–
18. Fabricated Structural Metal Products	–	–	–	–	–	–	0.3	–	–	–
19. Other Fabricated Metal Products	1.9	19.3	0.1	*	0.4	1.5	7.1	2.3	0.1	6.2
20. Agricultural, Mining and Construction Machinery	1.3	–	–	–	–	–	–	–	–	–
21. Metalworking Machinery	–	–	–	–	–	–	–	–	–	–
22. Other Machinery (except electric)	–	0.5	–	9.6	1.3	0.6	0.6	1.9	2.3	†
23. Motors and Generators	–	–	–	–	–	–	–	–	–	–
24. Radios	–	–	–	–	–	–	–	–	–	–
25. Other Electrical Machinery	–	–	–	–	–	–	–	–	–	†
26. Motor Vehicles	2.5	0.1	–	–	–	†	–	–	–	–
27. Other Transportation Equipment	0.2	–	–	–	–	–	–	*	–	†
28. Professional and Scientific Equipment	–	–	–	–	–	–	0.1	0.8	2.1	0.6
29. Miscellaneous Manufacturing	0.1	0.4	–	1.1	16.2	†	0.9	2.1	–	1.4
30. Coal, Gas, and Electric Power	1.4	6.9	†	28.8	2.3	1.0	1.0	17.0	1.9	9.0
31. Railroad Transportation	9.9	19.5	0.2	25.8	3.8	6.2	2.9	30.9	4.4	13.7
32. Ocean Transportation	1.6	4.5	†	3.6	0.7	0.4	*	2.2	*	2.1
33. Other Transportation	12.5	13.0	0.2	21.7	1.6	6.0	2.2	16.1	1.6	4.5
34. Trade	30.6	14.8	0.4	62.6	23.4	2.6	3.2	24.3	2.0	8.3
35. Communications	†	1.5	†	2.5	1.2	0.4	0.3	1.1	2.5	1.1
36. Finance and Insurance	5.4	5.1	†	5.5	1.5	3.3	1.0	2.5	1.5	0.9
37. Rental	53.9	3.2	–	6.9	6.1	0.8	0.9	3.6	4.0	1.6
38. Business Services	0.2	18.9	1.0	19.5	6.1	0.8	3.1	3.0	3.8	20.3
39. Personal and Repair Services	8.3	4.2	*	0.8	0.2	1.8	0.2	0.6	1.3	0.5
40. Medical, Educational and Nonprofit Organizations	–	–	–	–	–	–	–	–	–	–
41. Amusements	–	–	–	–	–	–	–	–	–	–
42. Scrap and Miscellaneous Industries	–	–	–	6.6	–	–	–	34.5	–	5.3
43. Undistributed	–	73.1	1.3	120.3	83.0	38.0	17.7	27.7	39.6	83.2
44. Eating and Drinking Places	–	–	–	–	–	–	–	–	0.1	–
45. New Construction and Maintenance	4.5	4.2	†	10.7	1.0	0.5	0.4	5.8	1.0	1.7
48. Government	18.3	40.3	1.0	175.6	23.8	14.6	6.0	68.5	21.8	36.5
50. Households[d]	431.7	222.3	3.8	902.8	254.3	110.6	57.3	297.9	197.2	164.1

Industry (Sector) Producing \ Industry (Sector) Purchasing	11. Products of Petroleum and Coal	12. Rubber Products	13. Leather and Leather Products	14. Stone, Clay, and Glass Products	15. Iron and Steel	16. Nonferrous Metals	17. Plumbing and Heating Supplies	18. Fabricated Structural Metal Products	19. Other Fabricated Metal Products	20. Agricultural, Mining and Construction Machinery	21. Metalworking Machinery	22. Other Machinery (except electric)	23. Motors and Generators	24. Radios	25. Other Electrical Machinery
1. Agriculture and Fisheries	–	–	14.3	*	–	1.0	–	–	–	–	–	–	–	–	–
2. Food and Kindred Products	*	–	129.7	0.1	0.1	*	–	–	*	–	–	*	–	–	*
3. Tobacco Manufactures	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
4. Textile Mill Products	†	69.2	25.7	1.6	–	–	*	*	1.0	†	0.2	2.1	0.1	0.8	5.
5. Apparel	–	–	0.6	0.1	–	–	*	–	0.1	*	–	0.2	–	*	*
6. Lumber and Wood Products	0.1	–	5.0	0.8	1.6	2.6	1.4	0.2	4.1	0.1	0.7	7.8	0.2	0.8	5.
7. Furniture and Fixtures	–	–	–	–	–	–	*	–	–	–	–	0.6	–	8.5	0.
8. Paper and Allied Products	1.4	3.1	15.8	8.8	*	*	0.5	0.2	9.6	†	0.3	4.6	0.3	1.4	6.
9. Printing and Publishing	–	–	–	–	–	–	0.1	–	1.3	†	0.2	0.6	–	0.1	0.
10. Chemicals	2.6	94.1	36.8	5.7	4.5	7.9	1.3	0.6	10.7	0.1	0.8	7.3	0.7	0.8	24.
11. Products of Petroleum and Coal	59.0	1.9	0.6	2.5	38.3	4.6	0.4	0.1	1.5	0.1	0.7	2.8	0.1	0.2	2.
12. Rubber Products	†	6.4	14.6	0.4	*	*	0.1	0.1	0.6	0.6	0.3	6.2	0.1	0.2	4.
13. Leather and Leather Products	–	–	302.9	–	–	–	*	*	0.1	–	1.8	0.8	*	*	0.
14. Stone, Clay and Glass Products	0.6	1.1	1.5	21.2	8.1	3.1	0.5	0.2	6.3	0.1	1.6	6.7	1.7	0.6	12.
15. Iron and Steel	0.1	2.2	0.3	1.1	180.1	3.1	10.4	17.0	167.1	4.5	23.6	108.9	8.6	1.1	27.
16. Nonferrous Metals	†	*	*	0.6	14.7	241.7	5.7	1.9	33.1	0.2	3.8	42.9	2.9	2.2	90.
17. Plumbing and Heating Supplies	–	–	–	–	0.7	–	2.4	1.3	0.9	–	–	4.1	0.1	–	4.
18. Fabricated Structural Metal Products	–	–	–	–	0.3	–	1.2	1.3	2.2	0.3	0.2	7.8	*	–	1.
19. Other Fabricated Metal Products	1.0	1.9	4.7	0.2	1.1	0.4	4.4	2.8	25.9	0.8	11.0	38.2	2.0	5.3	29.
20. Agricultural, Mining and Construction Machinery	–	–	–	0.2	0.7	0.7	–	0.7	–	1.0	3.3	11.2	–	–	–
21. Metalworking Machinery	–	–	–	–	0.4	0.7	0.4	0.3	6.4	0.4	9.4	10.5	0.5	0.5	2.
22. Other Machinery (except electric)	0.1	–	–	0.1	1.2	0.4	6.7	1.3	5.0	2.6	11.2	66.2	3.7	0.2	11.
23. Motors and Generators	–	–	–	–	–	–	2.1	0.2	0.1	0.3	4.8	37.1	1.2	0.4	7.
24. Radios	–	–	–	–	–	*	0.1	–	1.6	–	–	0.5	0.6	20.3	13.
25. Other Electrical Machinery	–	–	–	0.3	0.4	4.3	3.2	0.3	15.2	0.2	2.6	18.5	4.2	13.8	48.
26. Motor Vehicles	*	–	–	†	*	*	–	–	3.9	0.2	1.2	*	–	–	1.
27. Other Transportation Equipment	*	*	–	*	0.1	0.1	–	–	–	–	–	–	*	–	–
28. Professional and Scientific Equipment	–	–	–	†	*	–	2.1	0.1	0.4	†	0.3	4.2	0.1	0.3	1.
29. Miscellaneous Manufacturing	–	*	6.4	0.4	0.2	*	0.1	†	2.2	*	0.2	5.3	1.5	2.5	8.
30. Coal, Gas, and Electric Power	6.8	5.8	4.4	10.1	10.9	9.7	0.8	0.5	6.3	0.2	2.3	8.0	0.6	0.5	5.
31. Railroad Transportation	3.3	5.6	10.5	7.2	19.1	9.3	1.2	1.1	9.4	0.4	2.0	11.7	0.8	1.1	6.
32. Ocean Transportation	1.1	*	0.3	0.7	1.4	4.8	*	*	*	*	*	*	–	*	*
33. Other Transportation	5.7	1.1	6.1	3.5	6.3	1.8	0.4	0.2	2.1	0.1	0.5	3.4	0.2	0.7	2.
34. Trade	0.2	8.6	16.7	2.6	9.8	13.0	2.2	1.7	13.4	0.4	3.6	21.4	1.3	2.8	11.
35. Communications	0.2	0.9	1.5	0.5	0.7	0.6	0.2	0.2	1.6	0.1	0.7	2.7	0.2	0.3	1.
36. Finance and Insurance	1.5	1.1	2.0	2.3	2.0	1.3	0.4	0.4	2.3	0.1	1.2	3.7	0.5	0.8	3.
37. Rental	–	1.6	5.6	0.9	1.6	2.0	0.2	0.3	2.4	0.1	1.0	3.7	0.6	0.8	2.
38. Business Services	0.5	3.3	14.3	0.5	1.1	0.5	1.0	0.3	3.0	0.1	1.8	7.5	0.2	2.1	4.
39. Personal and Repair Services	0.2	0.2	0.3	1.5	0.1	0.3	0.1	0.1	0.4	†	0.2	0.6	*	*	0.
40. Medical, Educational and Nonprofit Organizations	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
41. Amusements	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
42. Scrap and Miscellaneous Industries	–	1.1	–	0.6	29.4	42.4	0.5	0.1	0.4	0.1	1.0	4.5	0.1	*	0.
43. Undistributed	9.6	51.2	94.4	28.1	13.0	9.4	12.8	5.3	121.6	3.0	41.2	171.4	3.3	17.0	113.
44. Eating and Drinking Places	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
45. New Construction and Maintenance	0.3	1.1	5.6	1.7	3.7	1.9	0.3	0.2	2.9	0.1	1.0	3.3	0.1	0.6	2.
48. Government	9.5	17.8	39.7	15.9	25.9	22.8	5.2	3.5	33.7	1.4	14.2	60.2	4.7	5.2	37.
50. Households[d]	59.9	159.5	333.0	111.3	178.4	141.3	37.9	28.9	284.0	10.0	153.2	508.3	39.0	49.6	288.

Here is one method: the spatial result of a non-spatial example. A MDS map of a

50 by 50 Inter-Industry Input - Output Table

2450 possible interactions.. Magnitude shown by width of line.



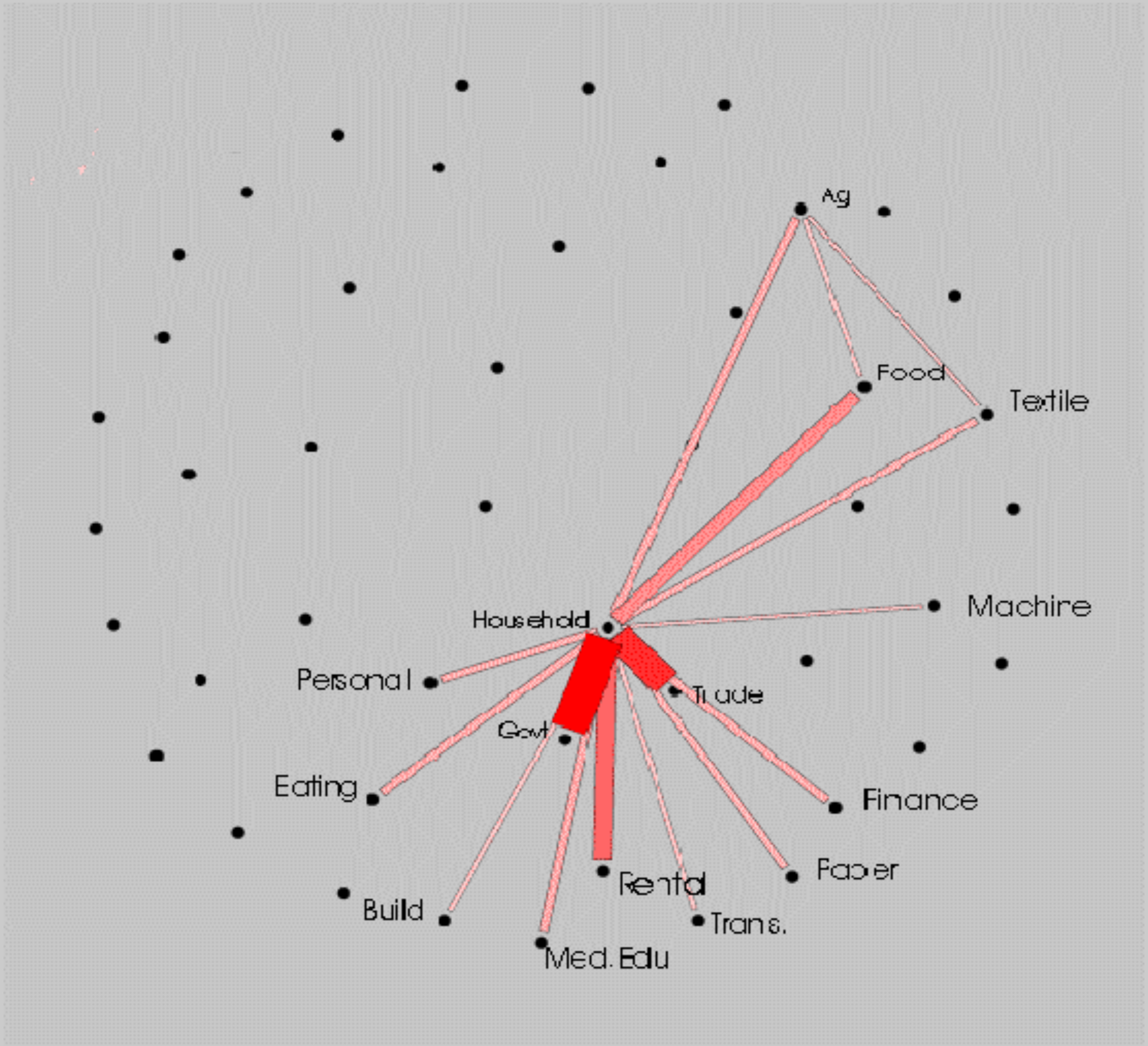
Would applying this to the Lieontiev inverse be of interest?

For more examples see www.geog.ucsb.edu/~tobler/publications/cartography/survey adjustment

From the Inter-Industry Input - Output table the MDS map shows that

15 of the 50 industries dominate

among 2450 possible interactions.



Iterative procedures, such as MDS, lend themselves particularly well to animation.

In the present instance two modes of animation are possible.

First, the rate of convergence of the discrepancy between the original guess and the current state can be shown. This should tend to a minimum function.

Secondly, and more interestingly, the simultaneous movement of the several objects between locations as they move towards equilibrium can make an amusing dynamic display.

Eighth Visualization

The **transform - solve - invert** paradigm

This is a classic way of solving problems.

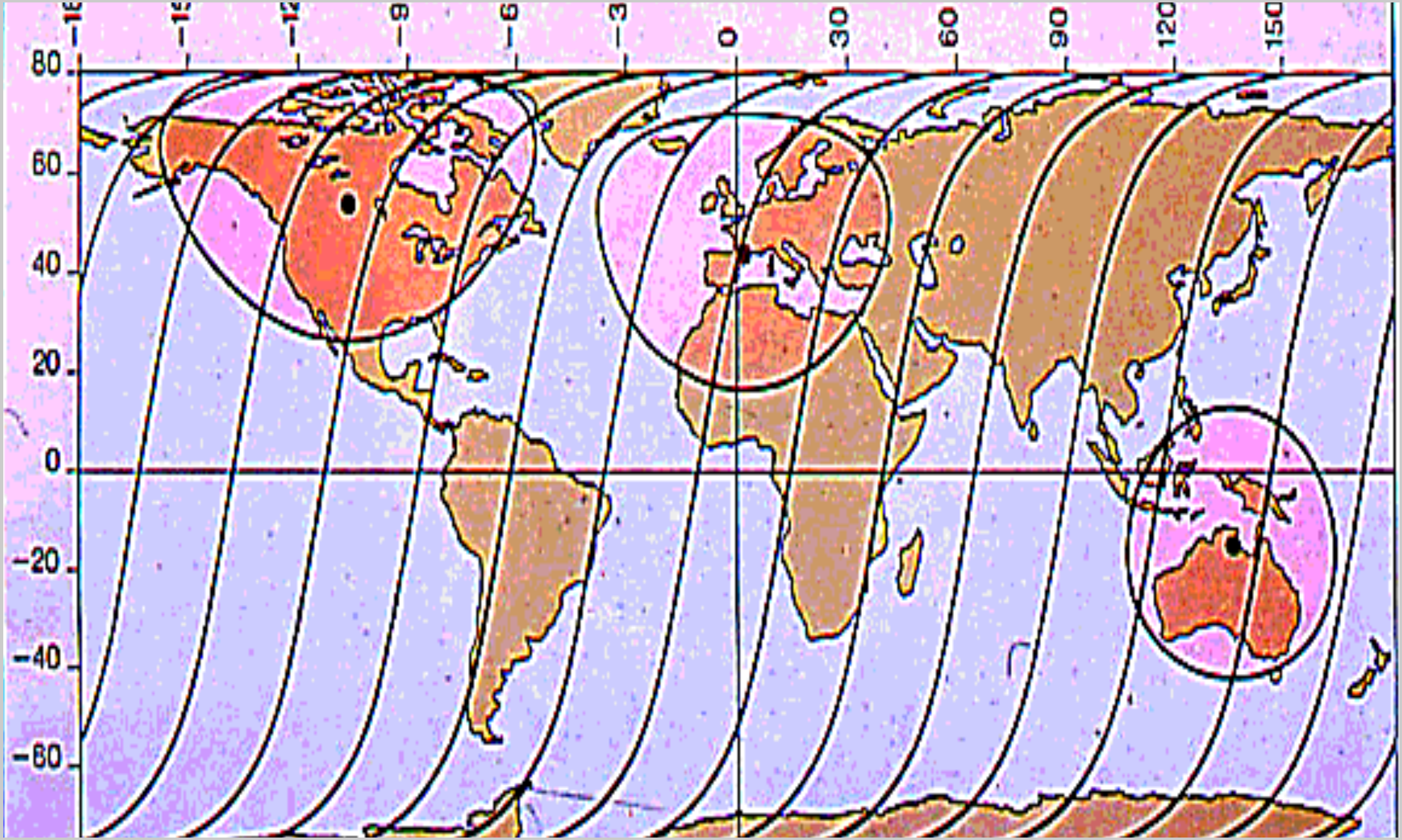
Change to a more appropriate coordinate system where the problem becomes simpler.

Solve the problem then revert to the original coordinates.

A map projection can be used in this manner.

Think of different map projections as different ways of producing different types of graph paper for spheres.

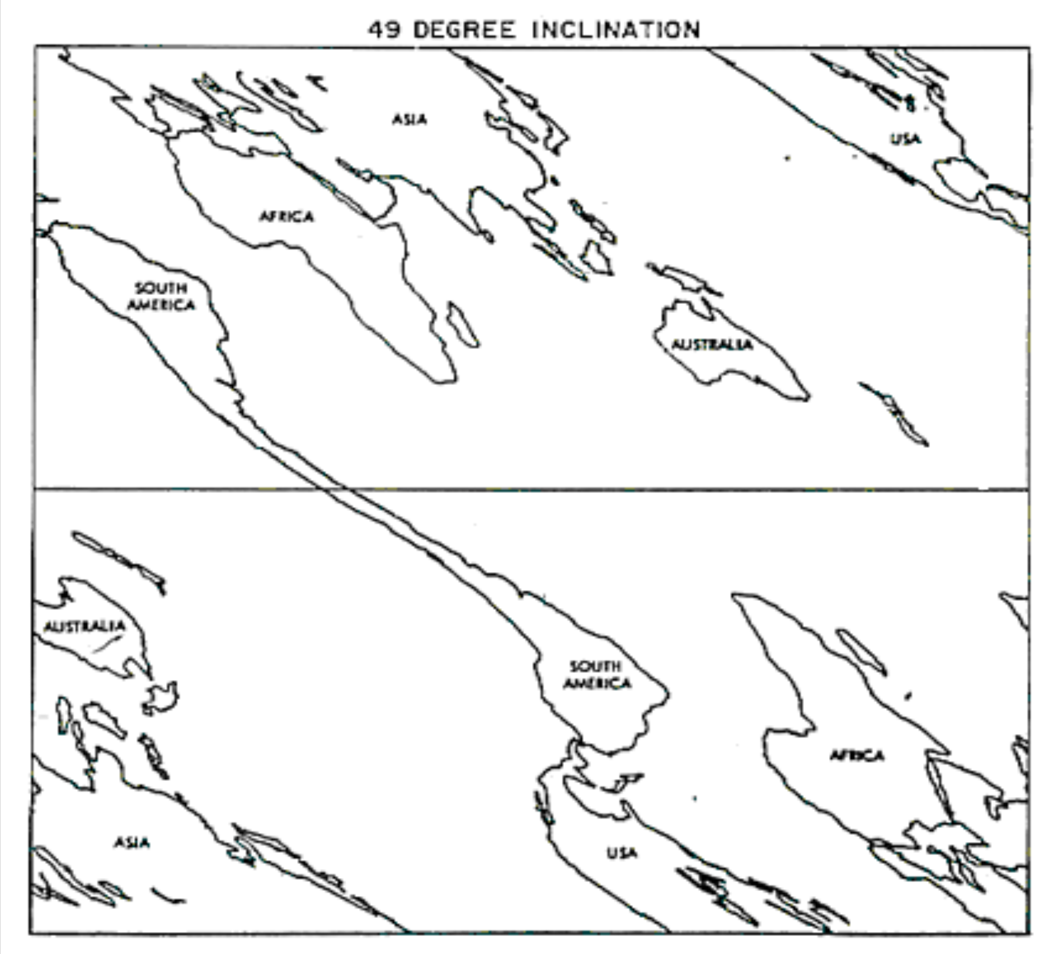
The conventional satellite tracking chart



The meridians and parallels are straight.

The satellite tracks are curves.

Breckman Chart to Track Satellites



Bend the meridians and parallels so that the satellite tracks are straight.

It is then easier to track the satellite.

Mercator’ s projection is a famous anamorphose.

It is designed to solve a navigation problem.

The seaman plots the course as a straight line on the distorted Mercator map and follows the indicated directions.

This same course is a curve on the earth but it appears as a straight line on the distorted map, and this is much easier to draw.

Other anamorphoses (map projections) can be used to solve other problems.

Map Projections and Cartograms

The equal area condition for a map projection in spherical and plane rectangular coordinates is

$$\frac{\partial x}{\partial \lambda}\frac{\partial y}{\partial \varphi} - \frac{\partial x}{\partial \varphi}\frac{\partial y}{\partial \lambda} = R^2 \cos(\varphi)$$

The condition equation for an areal cartogram is

$$\frac{\partial x}{\partial \lambda}\frac{\partial y}{\partial \varphi} - \frac{\partial x}{\partial \varphi}\frac{\partial y}{\partial \lambda} = R^2 D(\varphi, \lambda) \cos(\varphi)$$

Where $D(\phi, \lambda)$ is the density distribution on the earth, considered spherical.

Clearly, when the density distribution is constant, then the cartogram becomes an equal area map projection.

In both cases the one condition does not suffice to yield the two equations $x=f(\phi, \lambda)$, $y=g(\phi, \lambda)$ needed to completely define a map projection. The obvious second condition is to require that the angular distortion be minimized.

Another map projection example to solve another problem in the transform-solve-invert paradigm.

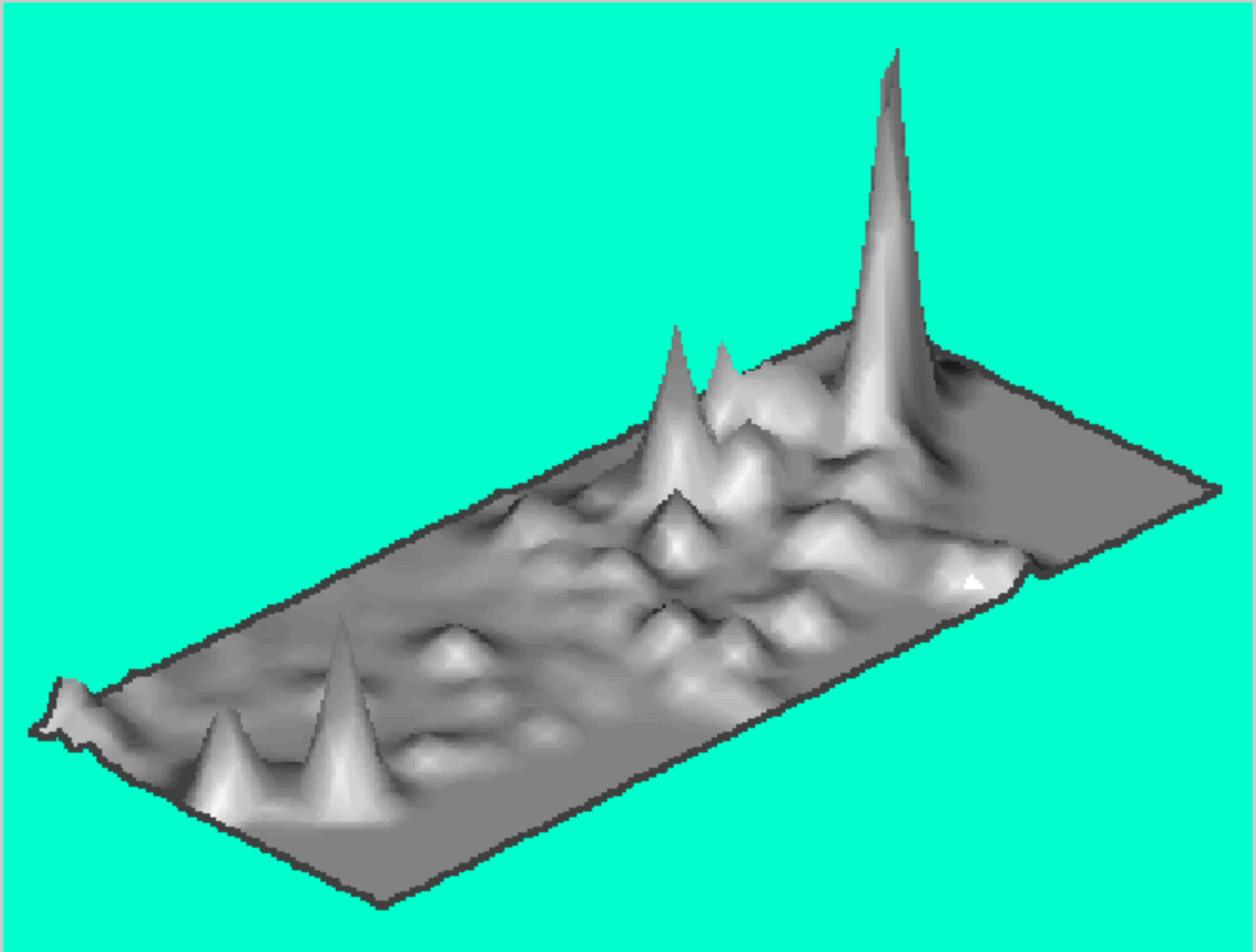
The next schematic illustrates the U.S. population density in perspective.

We would like to partition the U.S. into regions containing the same number of people.

There follows a map projection (cartogram) that may be useful for this purpose.

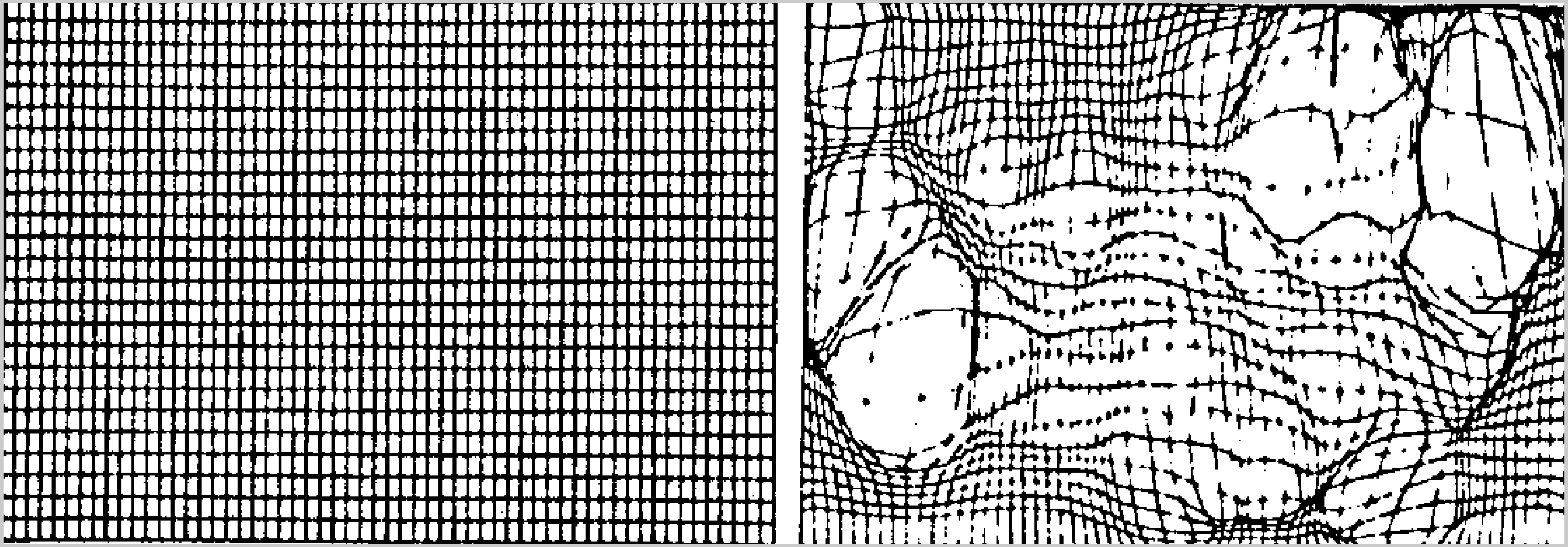
U.S. Population Density

by one degree latitude/longitude quadrangles



The Lat/Lon Grid in the Two Spaces

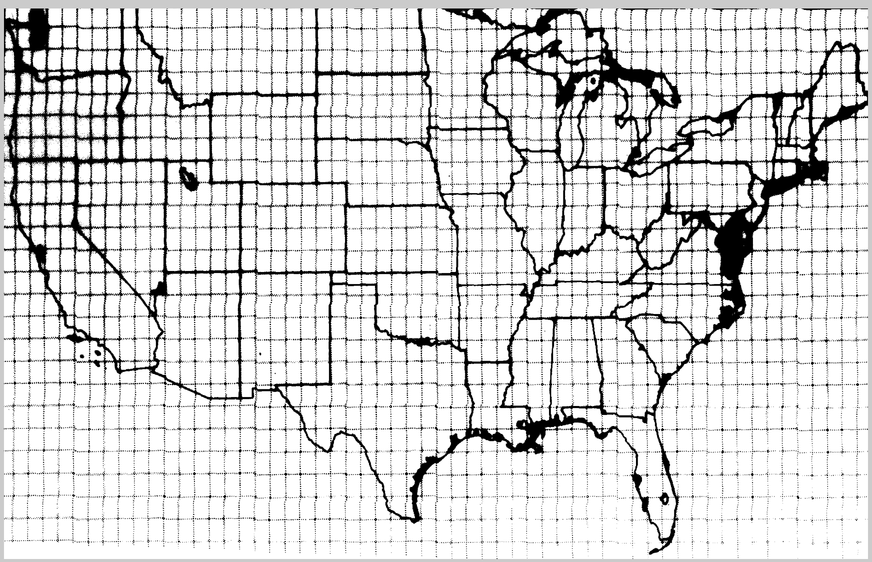
Left, the usual grid. Right, transformed according to population.

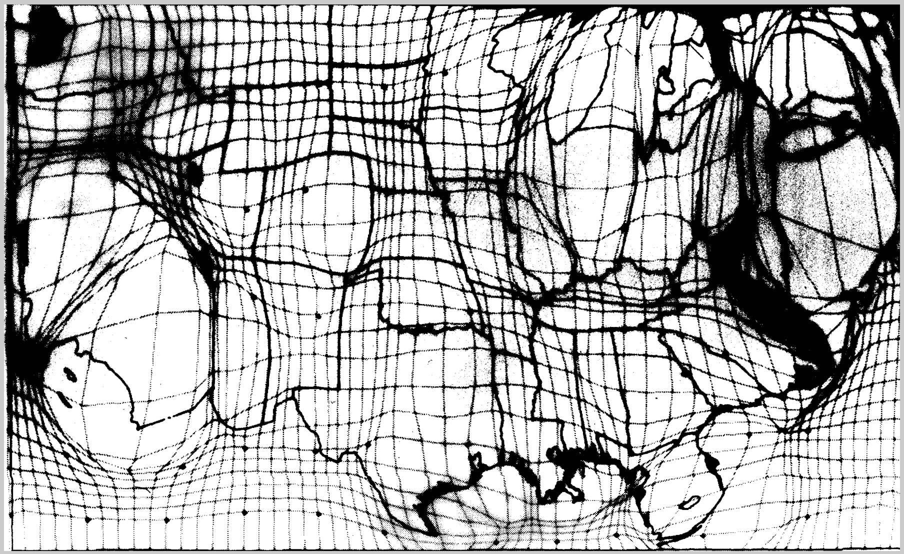


U.S. Map in the Two Spaces

Left, the usual map. Right, the transform.

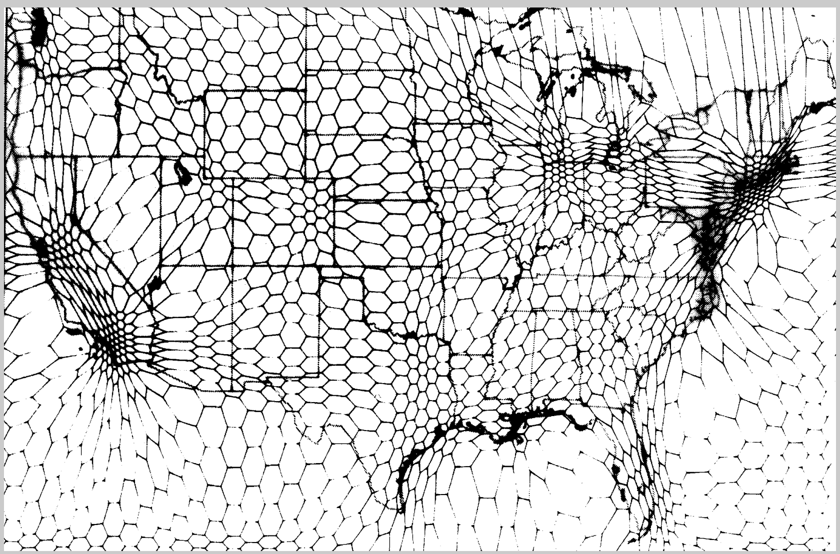
Map on the right adjusted to equalize the population.

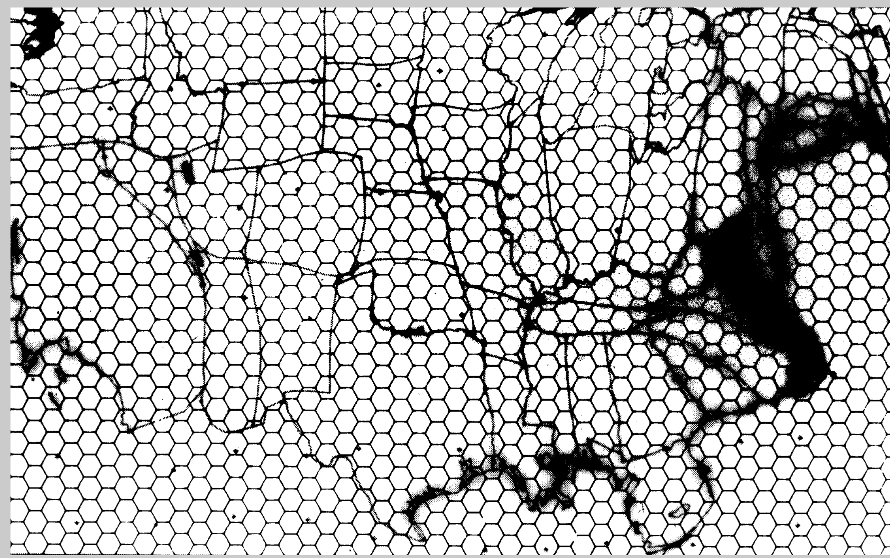




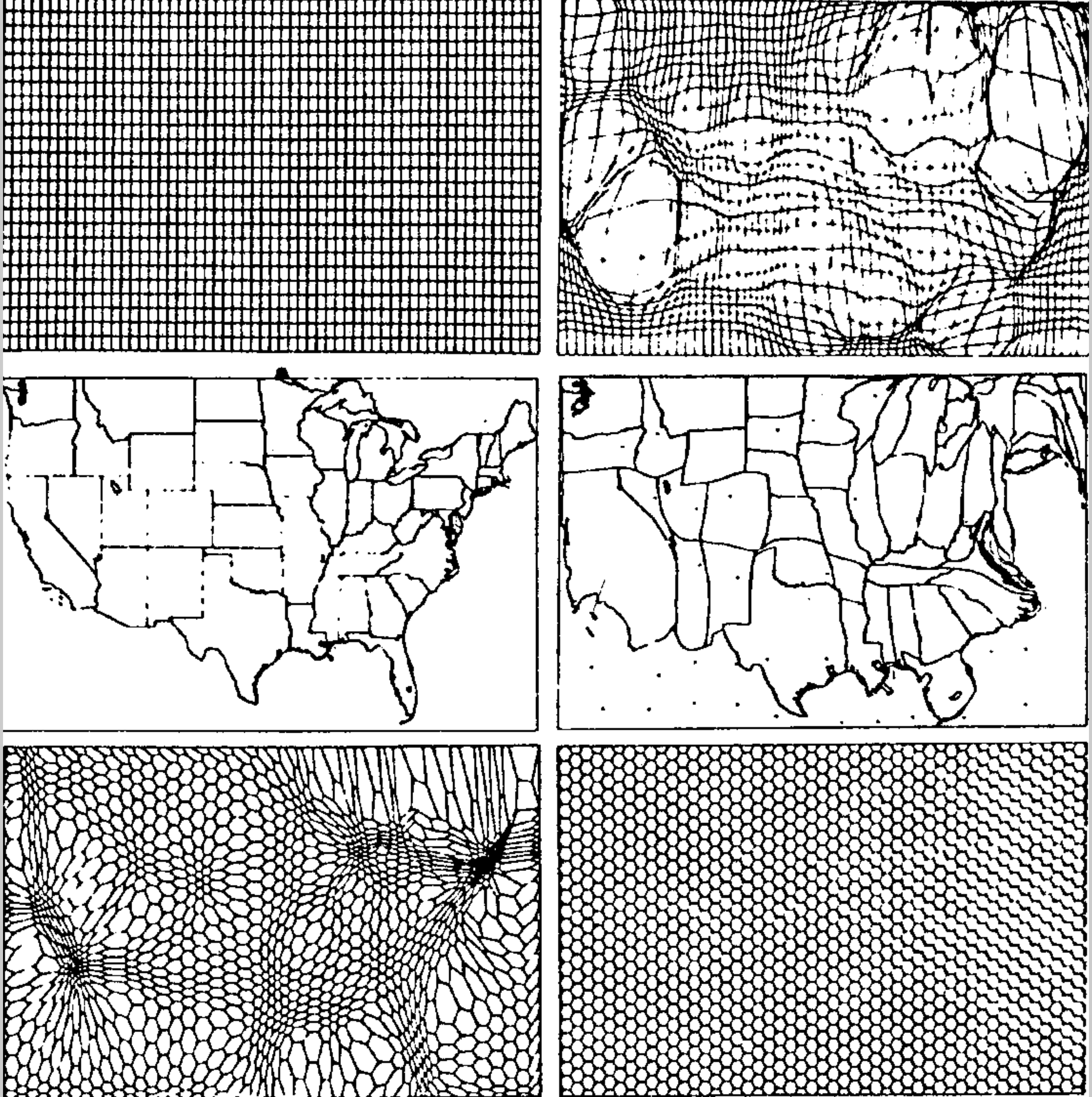
The Inversion

On the right: Uniform hexagons in the transformed space.
On the left: The solution.
The inverse transform partitions the U.S. into cells of approximately equal population.





W. Tobler, 1973, "A Continuous Transformation Useful for Districting", *Annals*, New York Academy of Sciences, 219: 215-220



Ninth Visualization

Map Matching

This example is from the field known as 'Mental Mapping'

A list of the sixty largest US cities, in alphabetical order, is given to students

Cities and Locations

Coordinates not given to students.

1	AKRON	41.066	-81.516
2	ALBUQUERQUE	35.083	-106.633
3	ATLANTA	35.749	-84.383
4	AUSTIN	30.299	-97.783
5	BALTIMORE	39.299	-76.633
6	BIRMINGHAM	33.499	-86.916
7	BOSTON	42.333	-71.083
8	BUFFALO	42.866	-78.916
9	CHARLOTTE	35.049	-80.833
10	CHICAGO	41.833	-87.749
11	CINCINNATI	39.166	-84.500
12	CLEVELAND	41.499	-81.683

Instructions to the Students

Work without any reference materials.

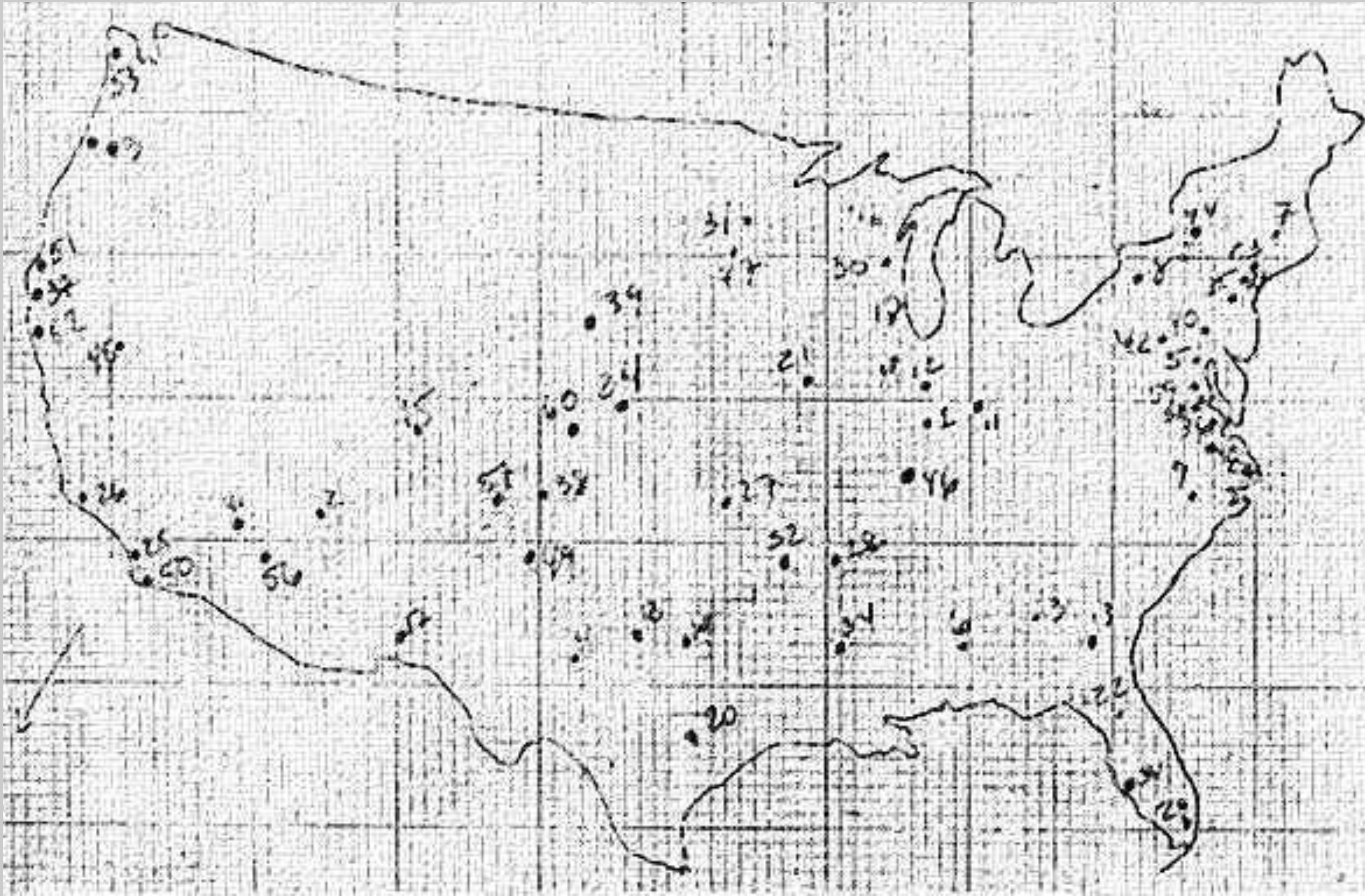
Use Graph Paper, wide Margin at top.

Plot Cities with ID Number on the Graph Paper.

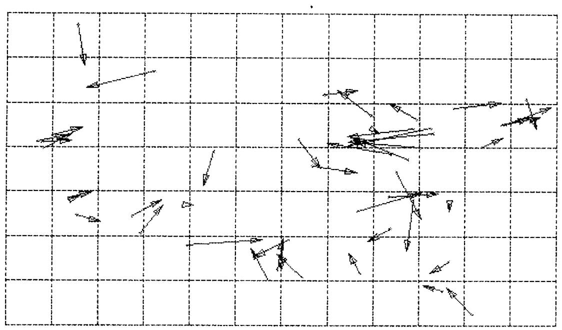
USA Outline may be drawn, but is not required.

An Anonymous Student's Map

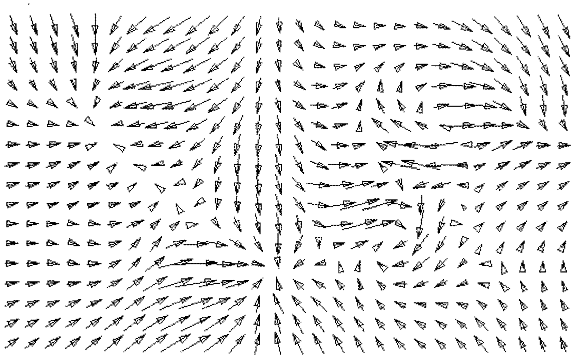
A very good student



Analysis of Student Data

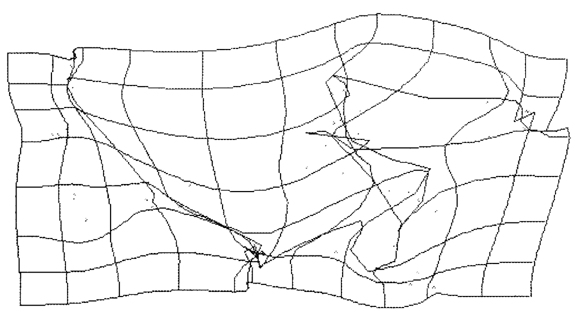


Displacement vectors



Interpolated vectors

Now Tissot’s indicatrix can be estimated and, by integration, a potential caculated.



Displaced grid

The Student Map

Shows displacement vectors

These vectors could also show change of address coordinates, due to a move, or commuting.

Or they could be home to shopping moves, etc.

Or they could be material moved or information and/or messages sent between places.

Thus there are many possible interpretations of this kind of vector displacement.

And there are methods of estimating the amount of strain and warping implied and calculation of an implicit potential field.

With student maps in hand

How to score?

Compute correlation, R^2, between actual and student estimates? How to do this?

Do you know how to compute correlation between fields of vectors?

Correlation between scores of different students? Factor analyze?

Compute vector field variance, etc., to determine degree of fuzziness?

Average vectors over all students?

I’m not going to answer these questions here, but

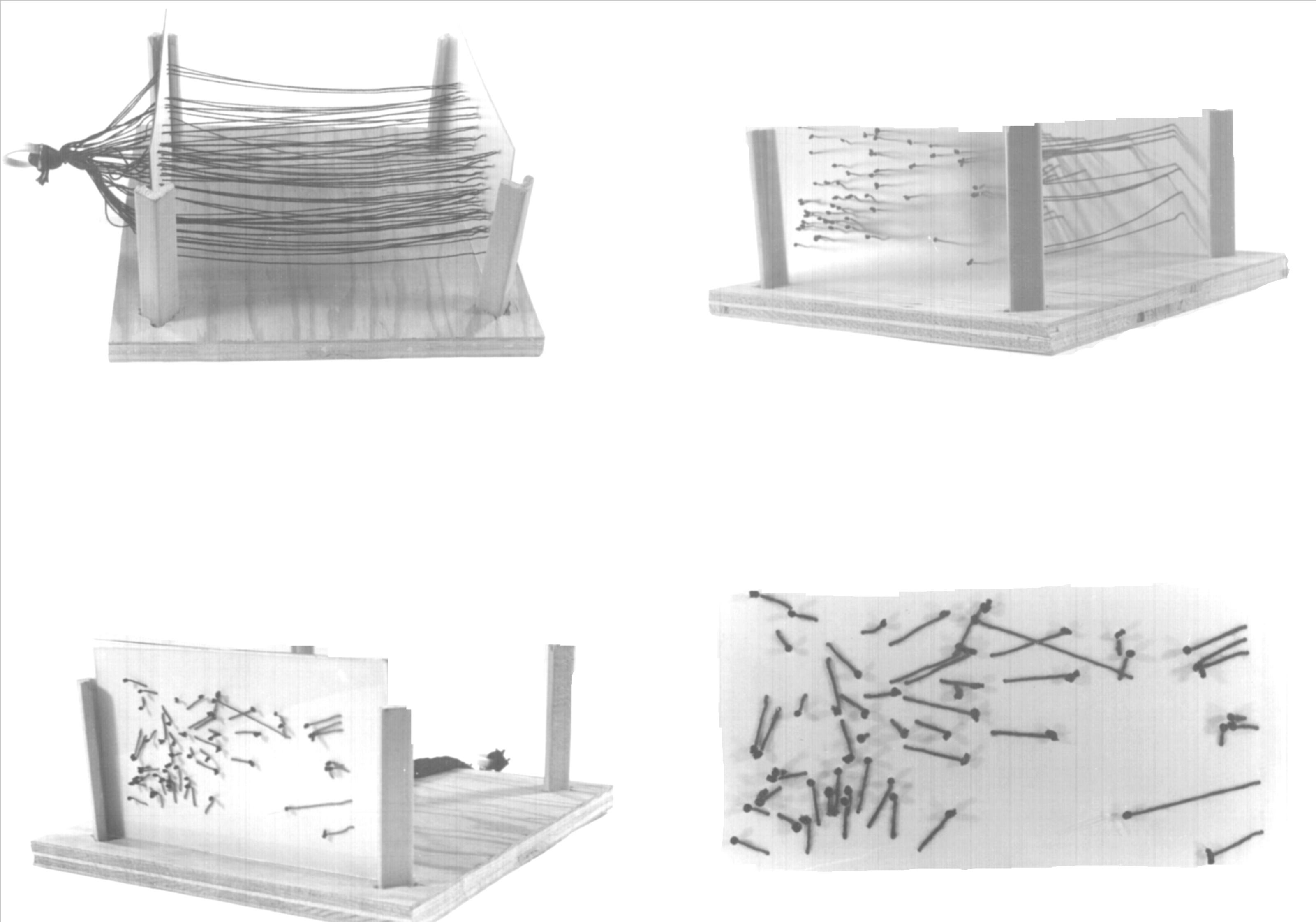
Is There a Method of Computing the Correlation Between Vector Fields?

The question comes up not only in meteorology and oceanography but also for the comparison of the student's maps, for comparison of old maps, and in many other situations. There are in fact such correlation methods, and associated with these are regression-like predictors. One of these is my bidimensional regression. Statistical significance tests are also available. For other approaches see, for example B. Hanson, et al, 1992, "Vector Correlation", *Annals*, AAG, 82(1):103-116.

To illustrate the scoring concept for students I have built

The Map Machine

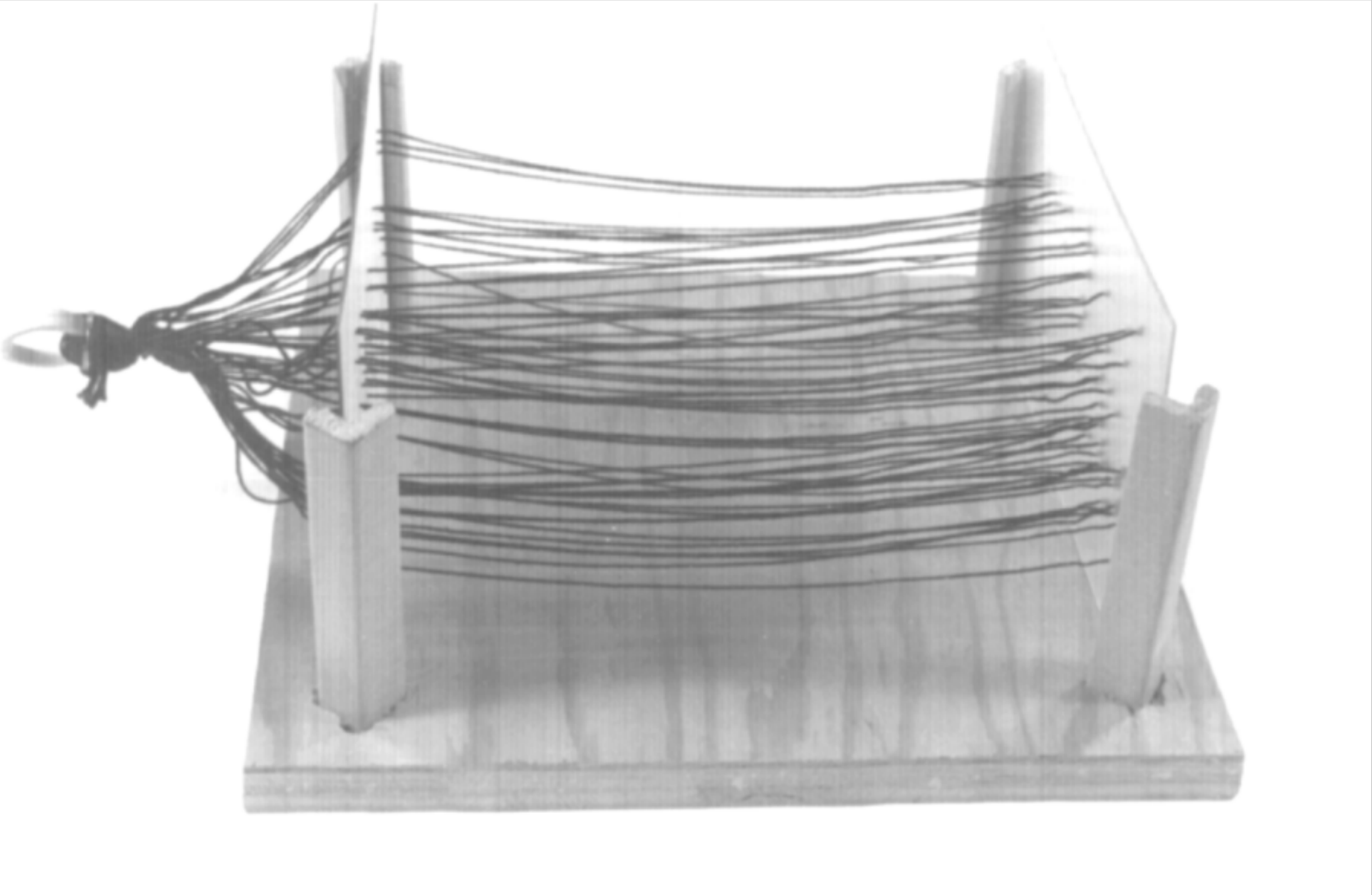
Showing the 1 to 1 correspondence of the locations & pulling them together



The Map Machine

Detail View 1

Showing the one to one correspondence between the images

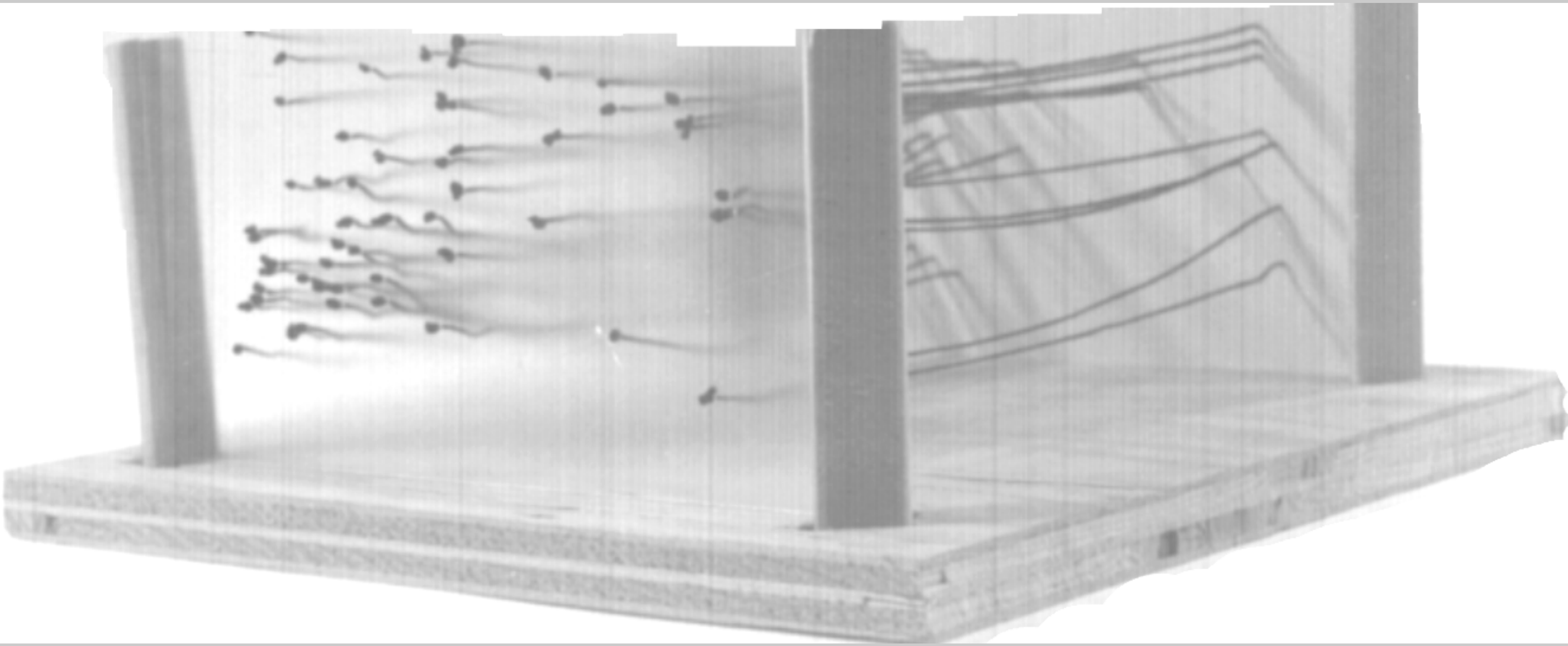


The Map Machine

Detail View 2

The front panel is transparent, back panel is white, strings are black

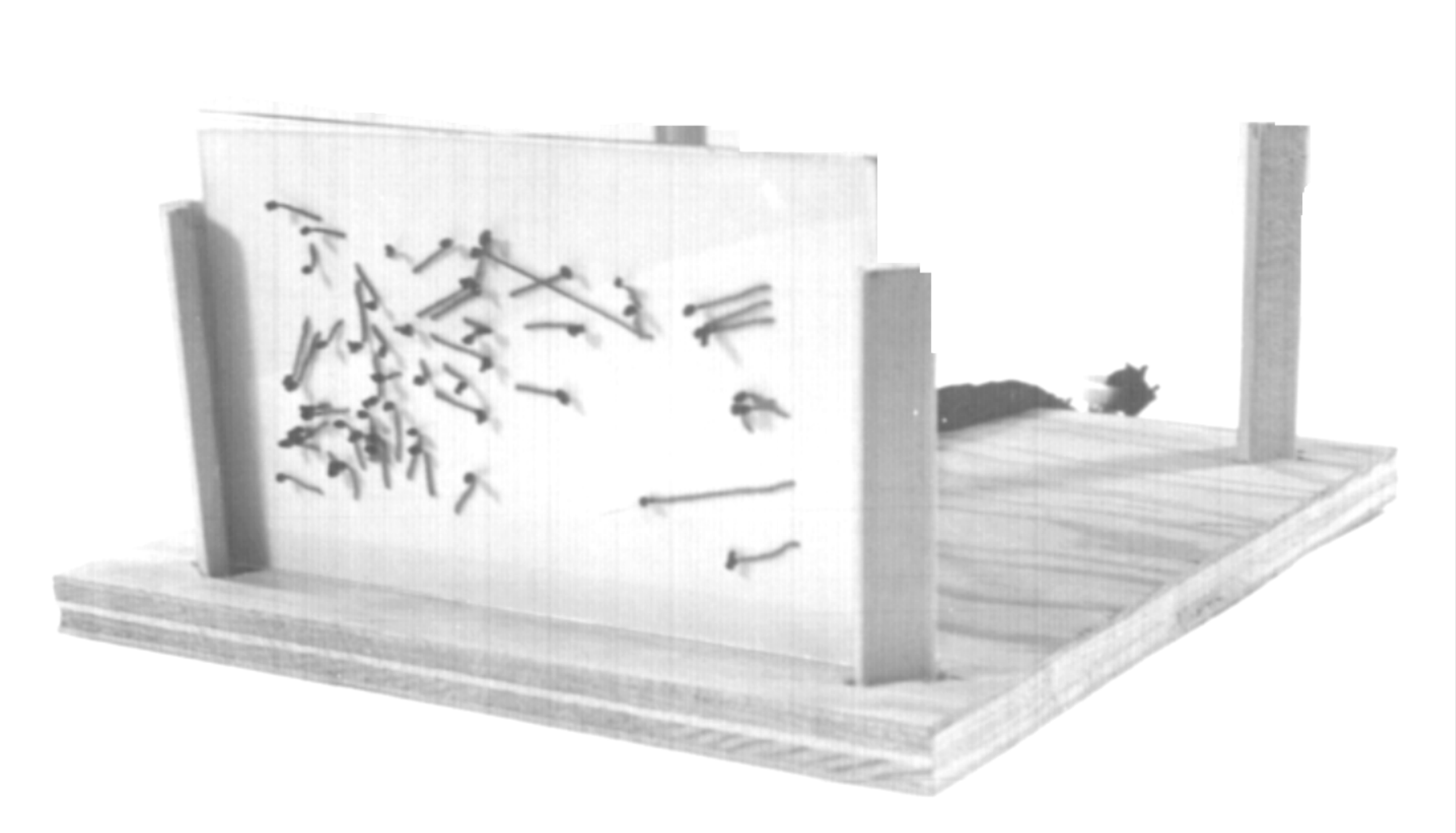
One set of locations is ‘correct’, the other is the estimate.



The Map Machine

Detail View 3

Releasing the back panel and pulling the strings together

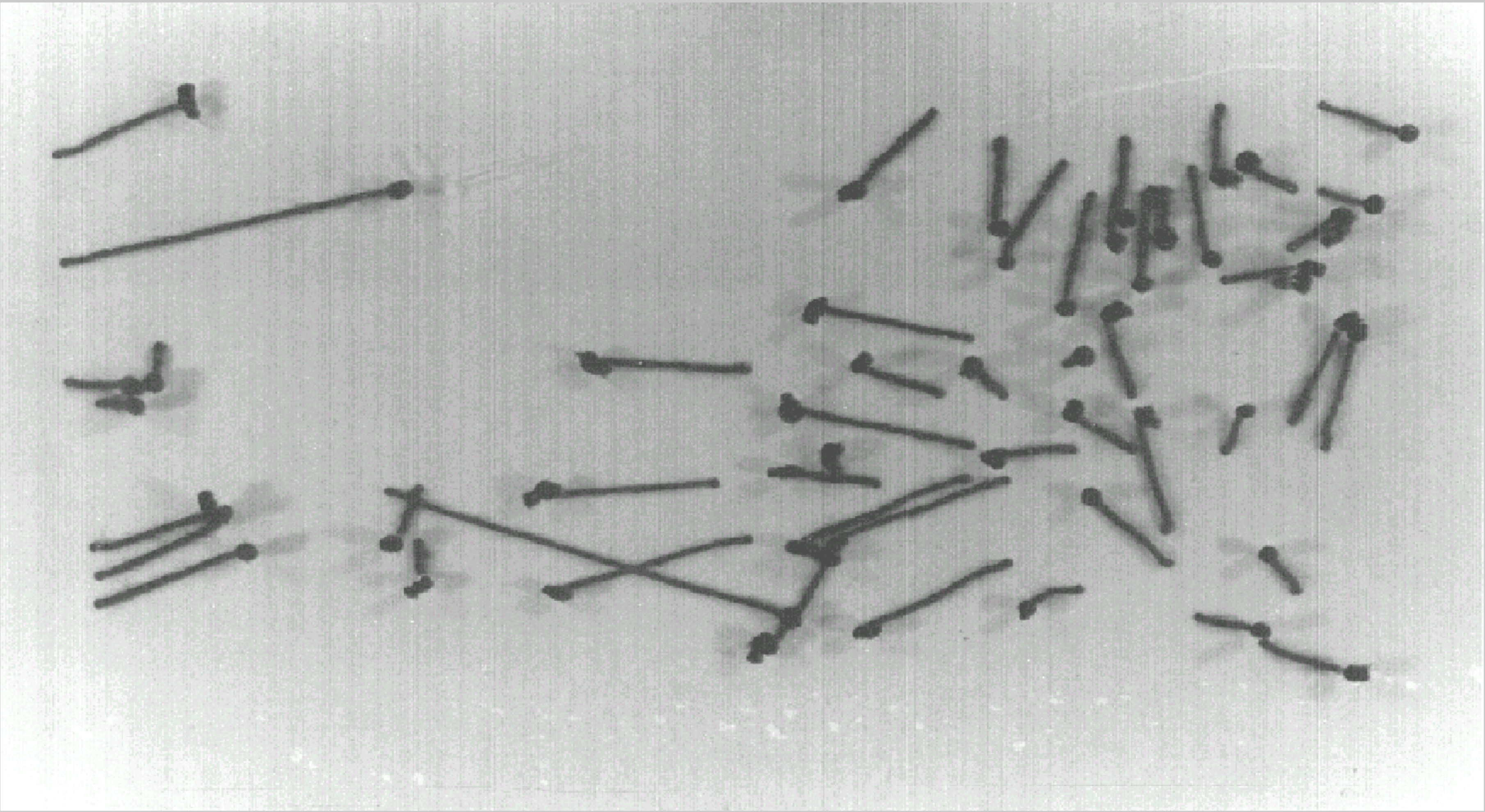


The Map Machine

The Final View

Corresponds to the computer image of displacements

Connecting the estimate to the ‘correct’ locations



My Bidimensional Regression Computer Program

Now available for download online

Differs from ordinary regression in that both the independent and the dependent variable each have two components.

One obtains the same effect as the map machine by estimating the difference between original and image components.

The original can be a modern map and the image an old map.

Or depiction of many kinds movements.

The difference between pairs of locations is used to compute displacements.

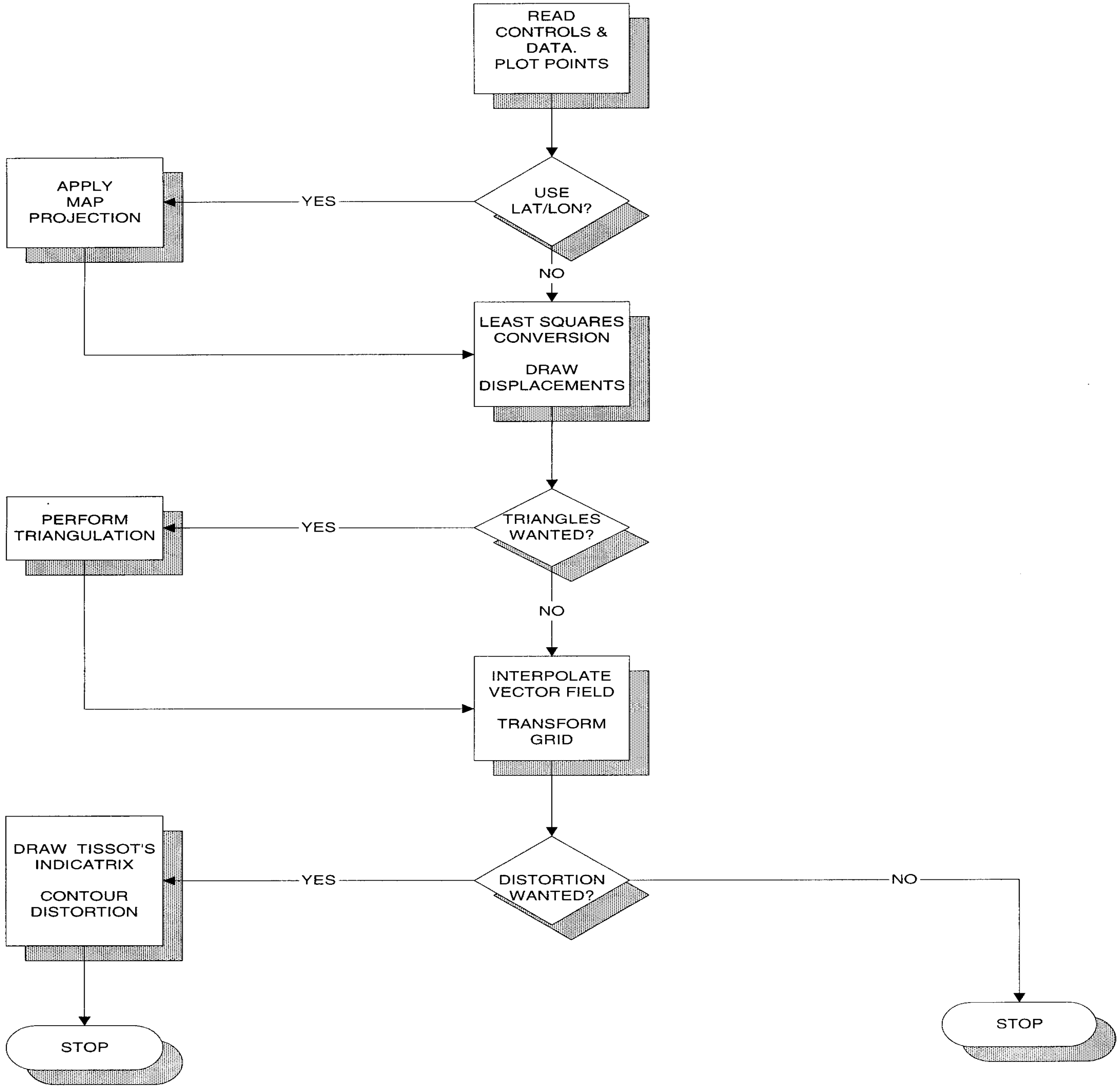
An interpolation can then performed to stretch one space to fit the other.

This is also known as ‘rubber sheeting’

The stretch can then examined mathematically,

and the distortion calculated.

BIDIMENSIONAL REGRESSION PROGRAM

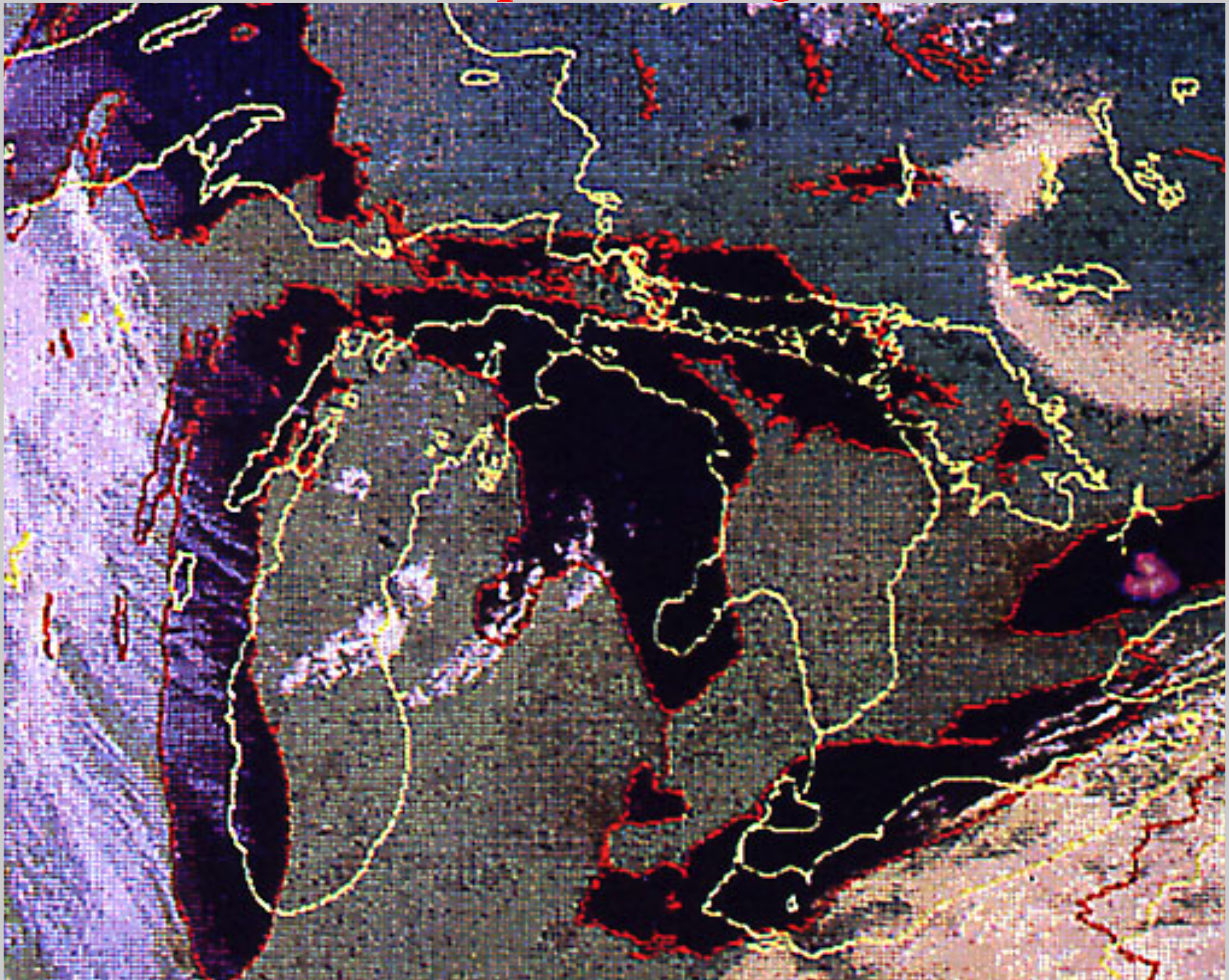


Waldo Tobler

Vectors Appear in Map Matching.

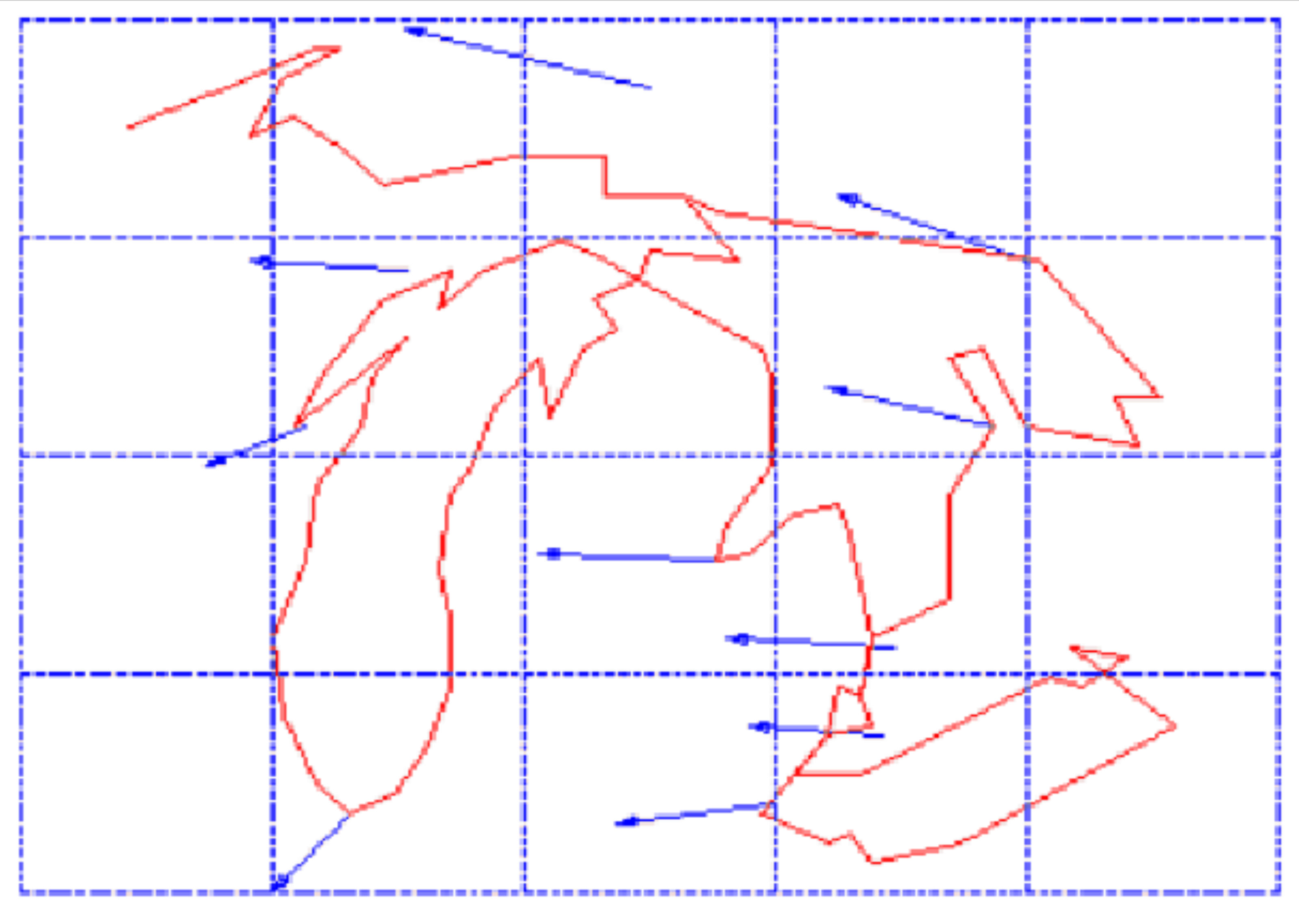
Here is an example

Map and Image



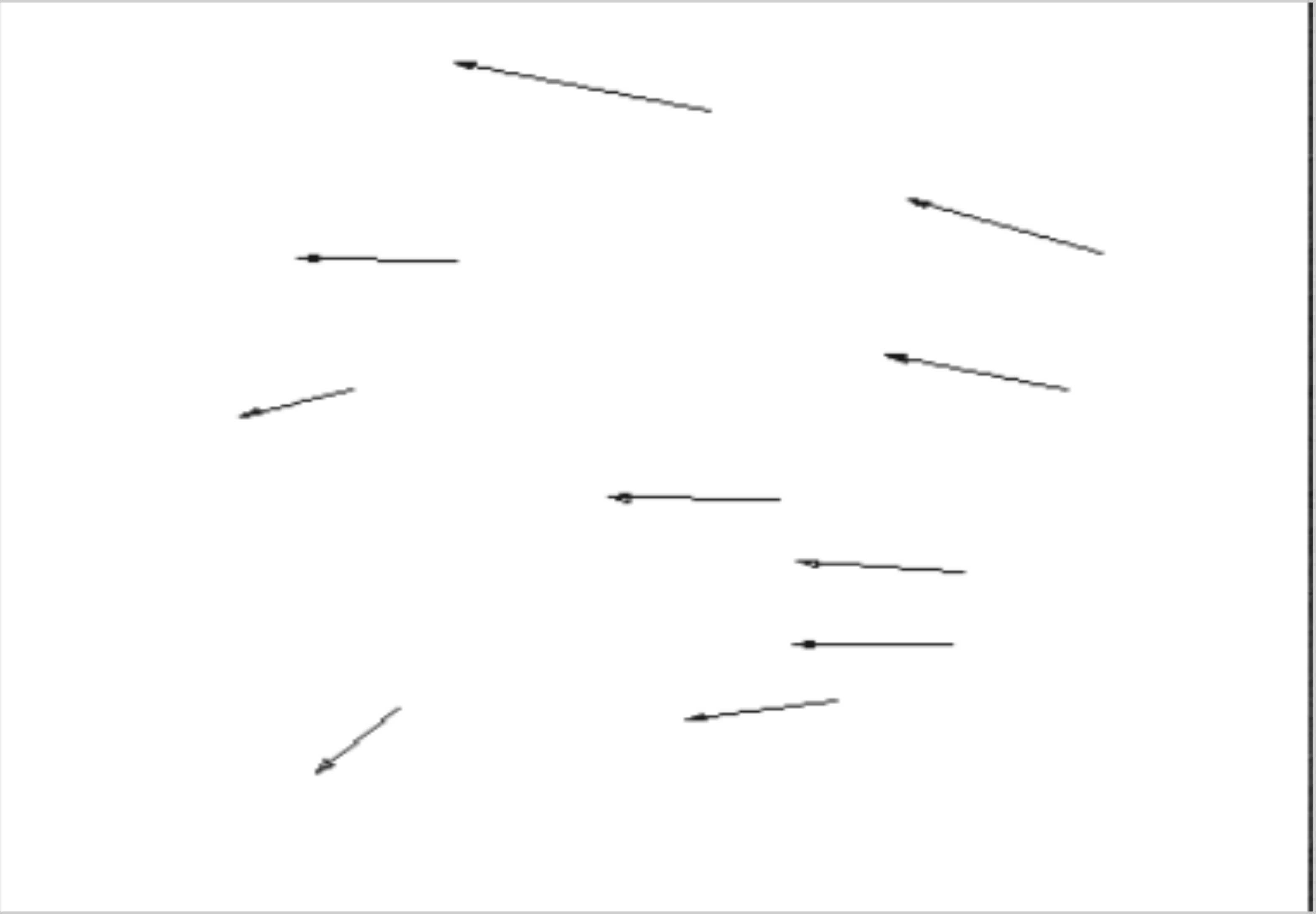
The difference between the map and the image

shown as discrete vectors

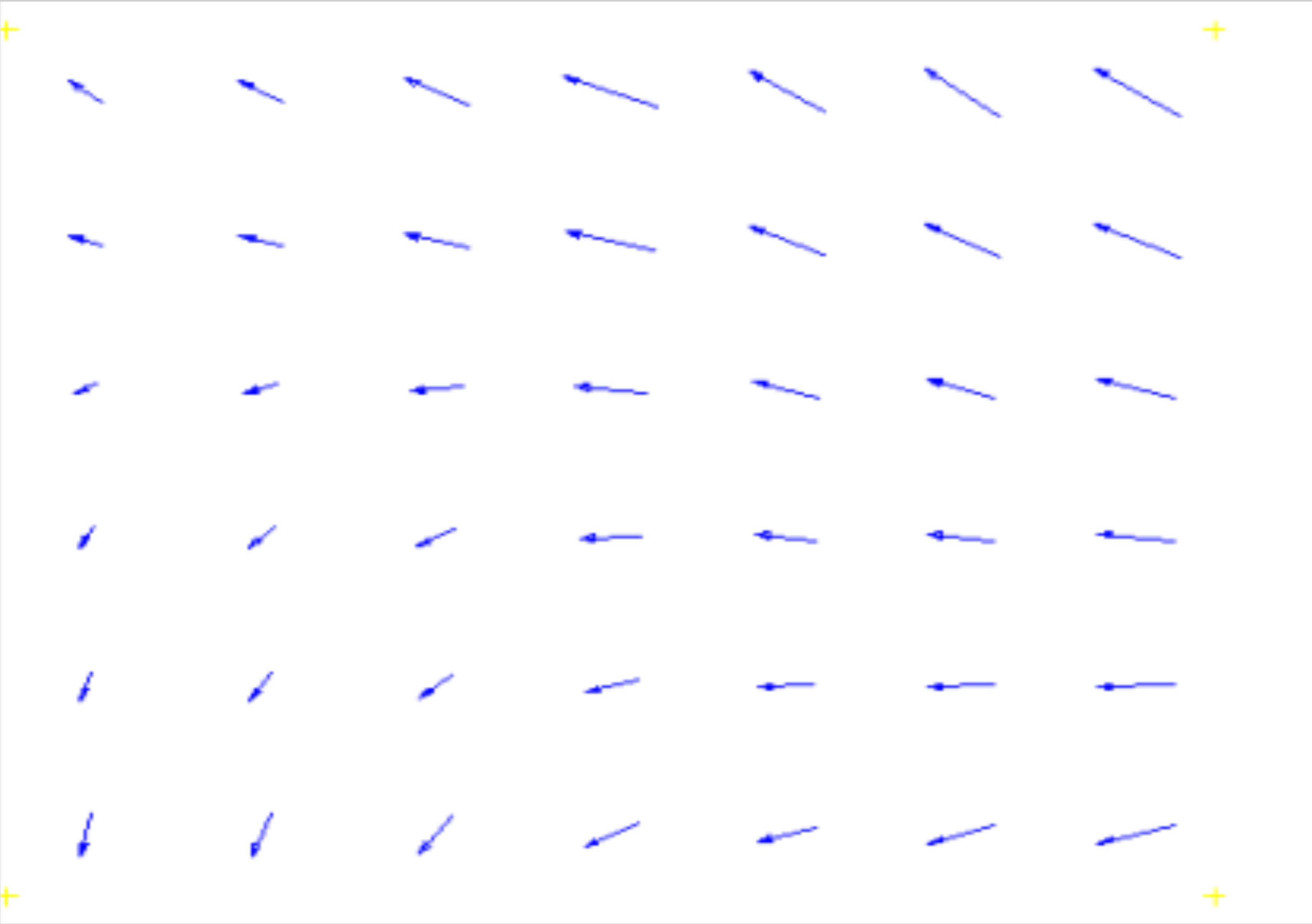


Difference Vectors by themselves, without the grid.

These can be interpolated to yield a vector field

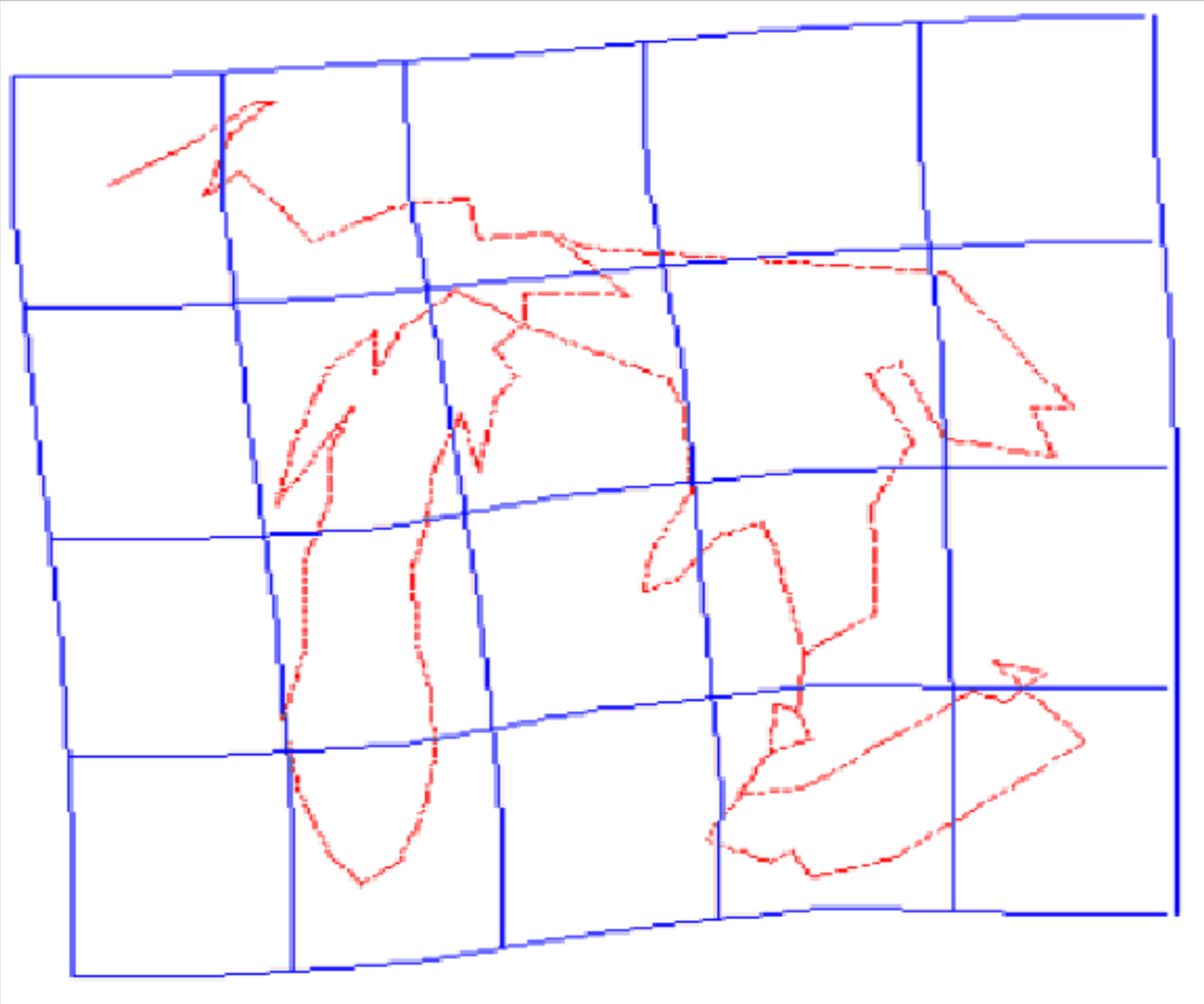


Interpolated Vector Field



Great Lakes Displaced

The grid has been ‘pushed’ by the interpolated vector field

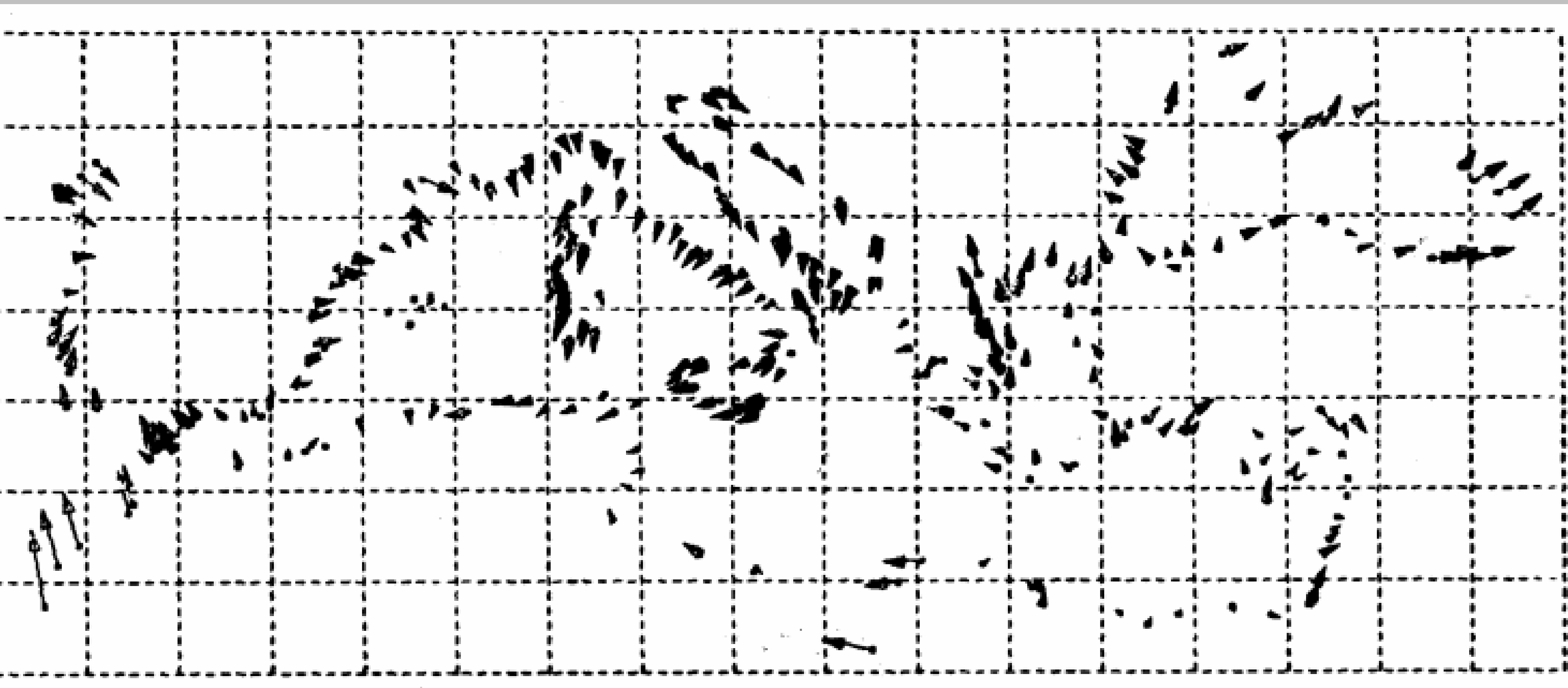


Benincasa Portolan Chart

Comparing positions on an old map with actual ones yields vectors.

Mediterranean Chart Displacements

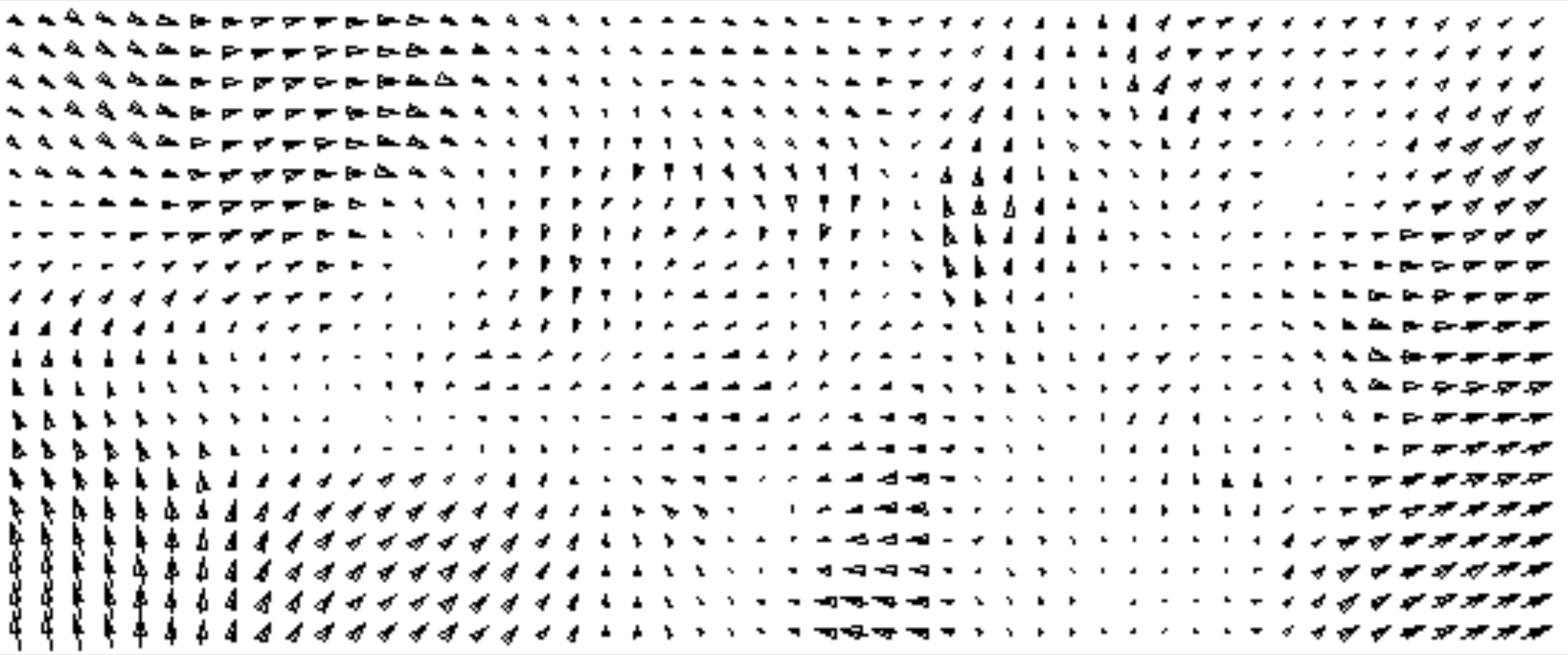
After Loomer



Interpolated Vector Field

Based on Mediterranean displacements from Loomer's thesis

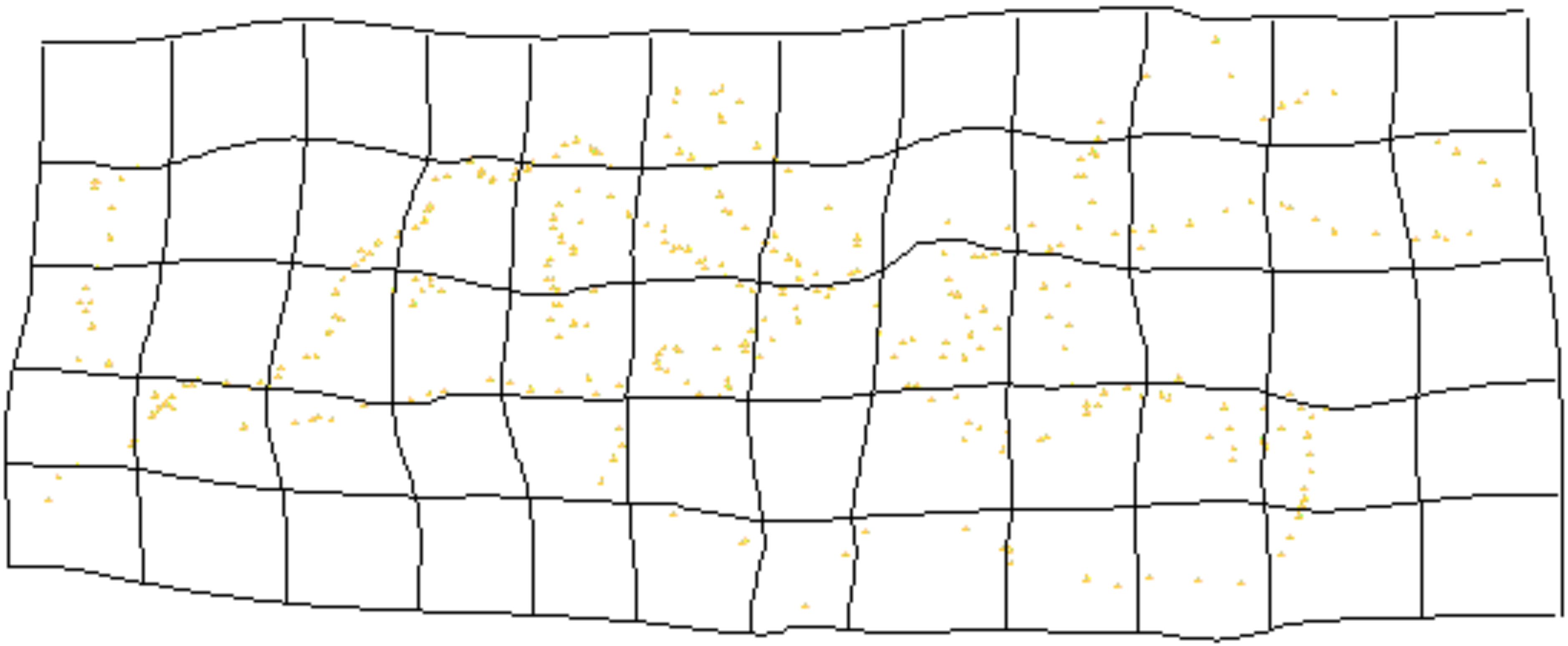
Computed and drawn at half scale using the Bidimensional Regression computer program



Warped Grid of the Portolan Chart

As 'pushed' by the interpolated vector field

Computed and drawn using the Bidimensional Regression computer program



Distortion of chart contoured.

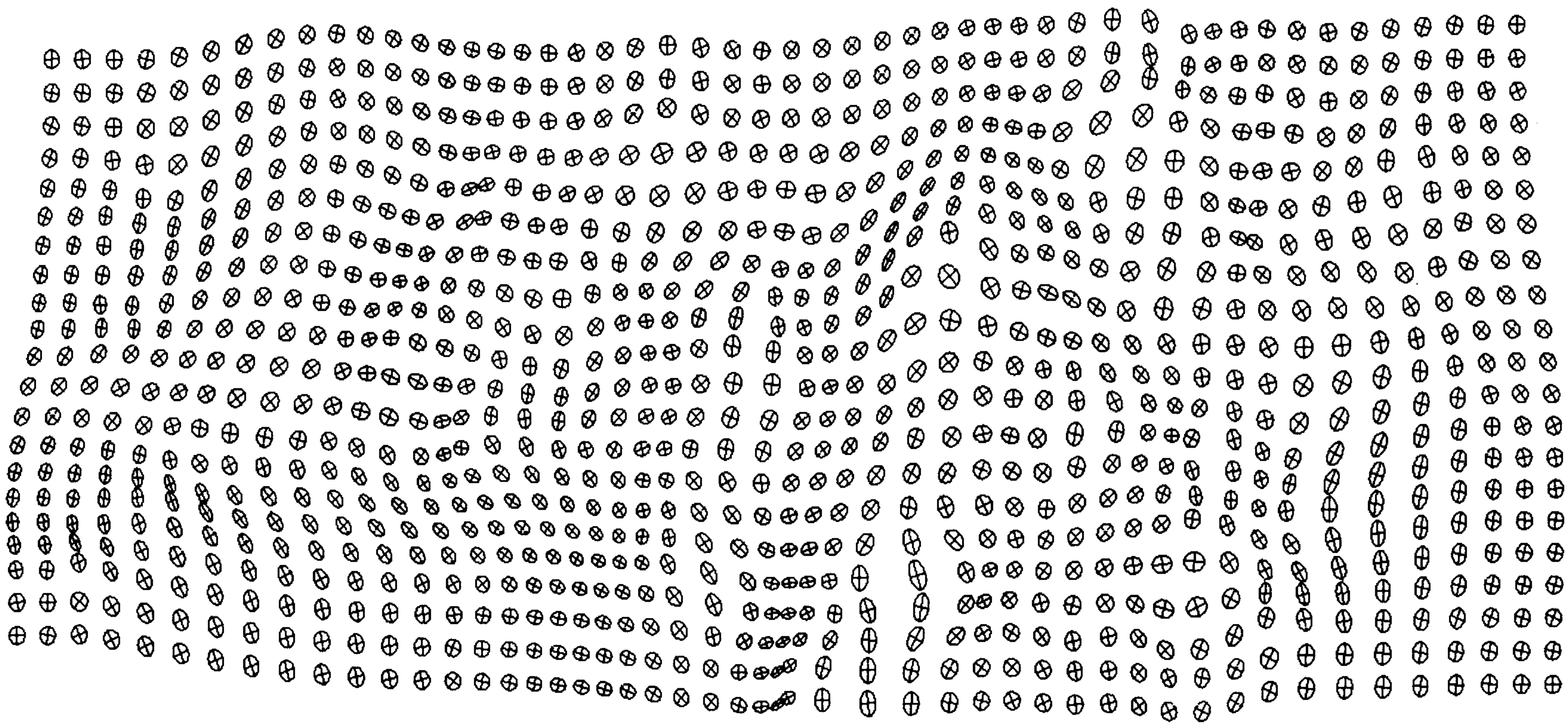
Using the sum of squares of the partial derivatives of the displacement vectors from the image.
Computed and drawn by the Bidimensional Regression program.

Tissot's measure of distortion for the chart.

Small circles in the original become ellipses in the transform.

Obtained from the displacement vectors and illustrating distance, angle, and area distortion at each location.

Calculated and drawn using the Bidimensional Regression program.



Tenth Visualization

Mass Preserving Reallocation using Areal Data.

Given observations for bounded areas,
for example within census tracts or school districts,
that is, **not at point locations or centroids**,
produce a quasi-continuous field on a raster for contouring,
or for data conversion between these or other areal units.

Both of these uses are considered important for regional studies.

For example, area boundaries change frequently, or zones do not coincide, requiring interpolation.

Begin by covering the polygons by a raster.
Then smooth the data transition between the polygonal units
but <u>keep</u> <u>the</u> <u>individual</u> <u>volumes</u> <u>intact.</u>

W. Tobler, 1979, "Smooth Pycnophylactic Interpolation for Geographic Regions", J. American Statistical Association, 74(367): 519-536.

Mathematically this requires the constrained minimum of an integral.

In other words, the objective requires that the total content within each region must remain its value, or

$\iint_i f(x, y)\, dx\, dy = V_i$ for each region i.

The smoothness requirement is given by the LaPlace equation

$$\nabla^2 x + \nabla^2 y = 0.$$

This says that the neighboring locations have similar values - or, in a raster, that the central value is the average of those surrounding it, or that the partial derivatives change slowly.

Recognize that this is a hypothesis about the phenomena under study.

But it immediately yields a computational algorithm.

What the Mathematics Means

Imagine that each unit is built up of colored clay, with a different color for each unit.

The volume of clay represents the number of people, say, and the height represents the density.

In order to obtain smooth densities a modeling spatula is used, but no clay is allowed to move from one unit into another.

Color mixing is not allowed.

The smoothing is done using an iterative process.

The first step is to “rasterize” the region.

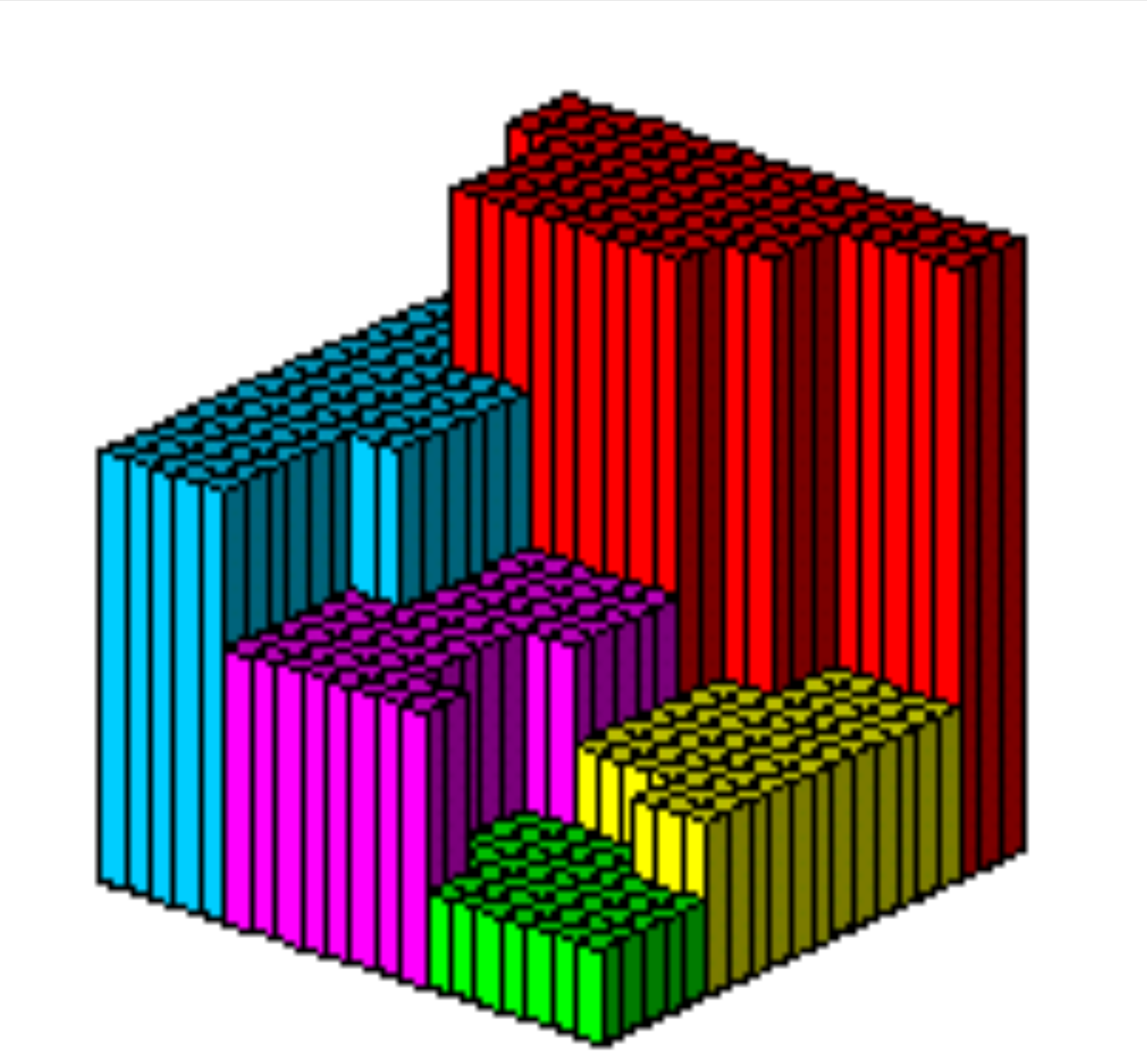
Then the smoothing is done on this raster, all the while maintaining the “population”.

The number of iteration steps depends on the size of the largest region, in raster units.

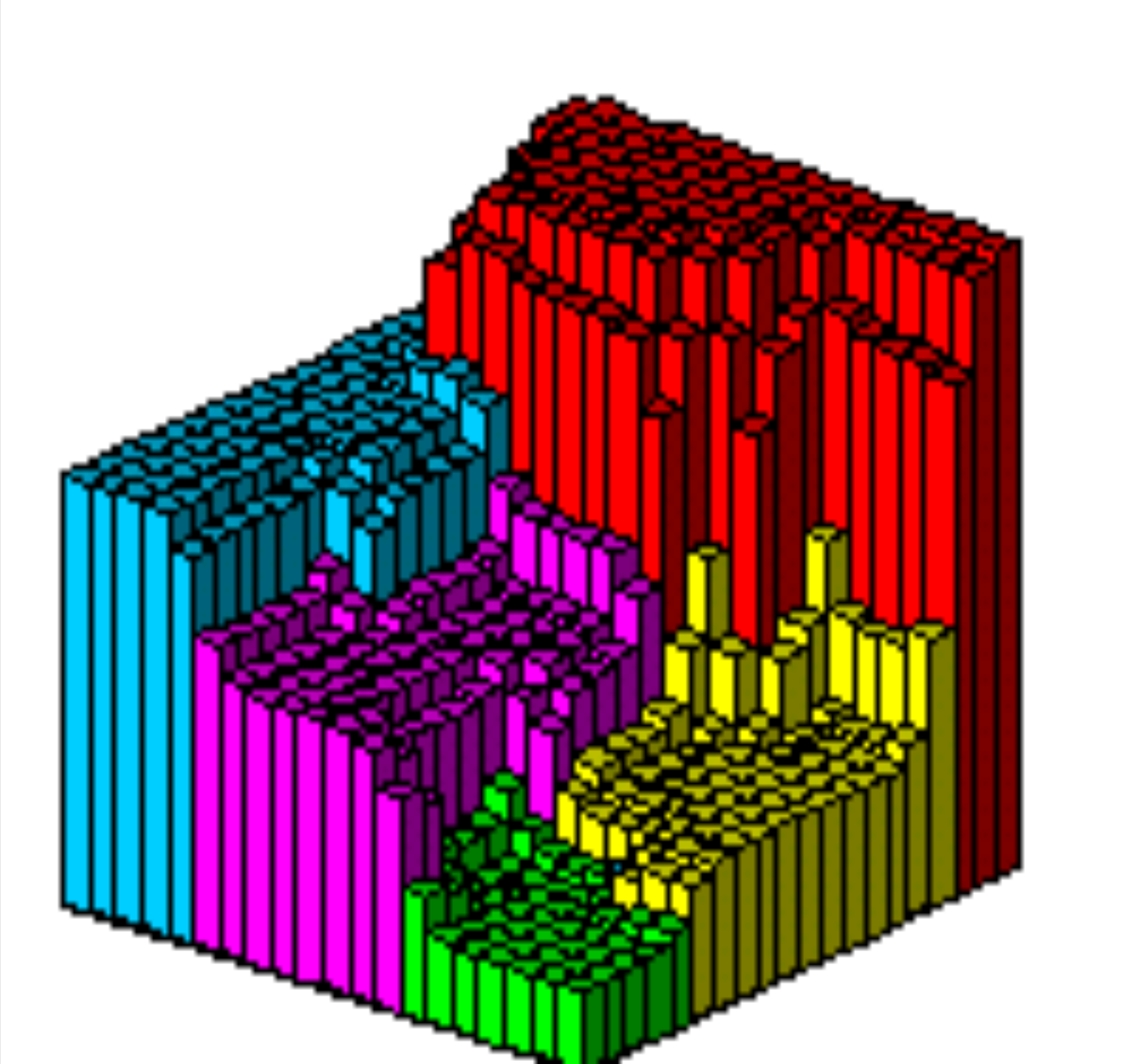
That is because the smoothing must cross from edge to edge of the largest region. The finer the raster, the higher the resolution, and the longer the iteration time.

Zero Iterations

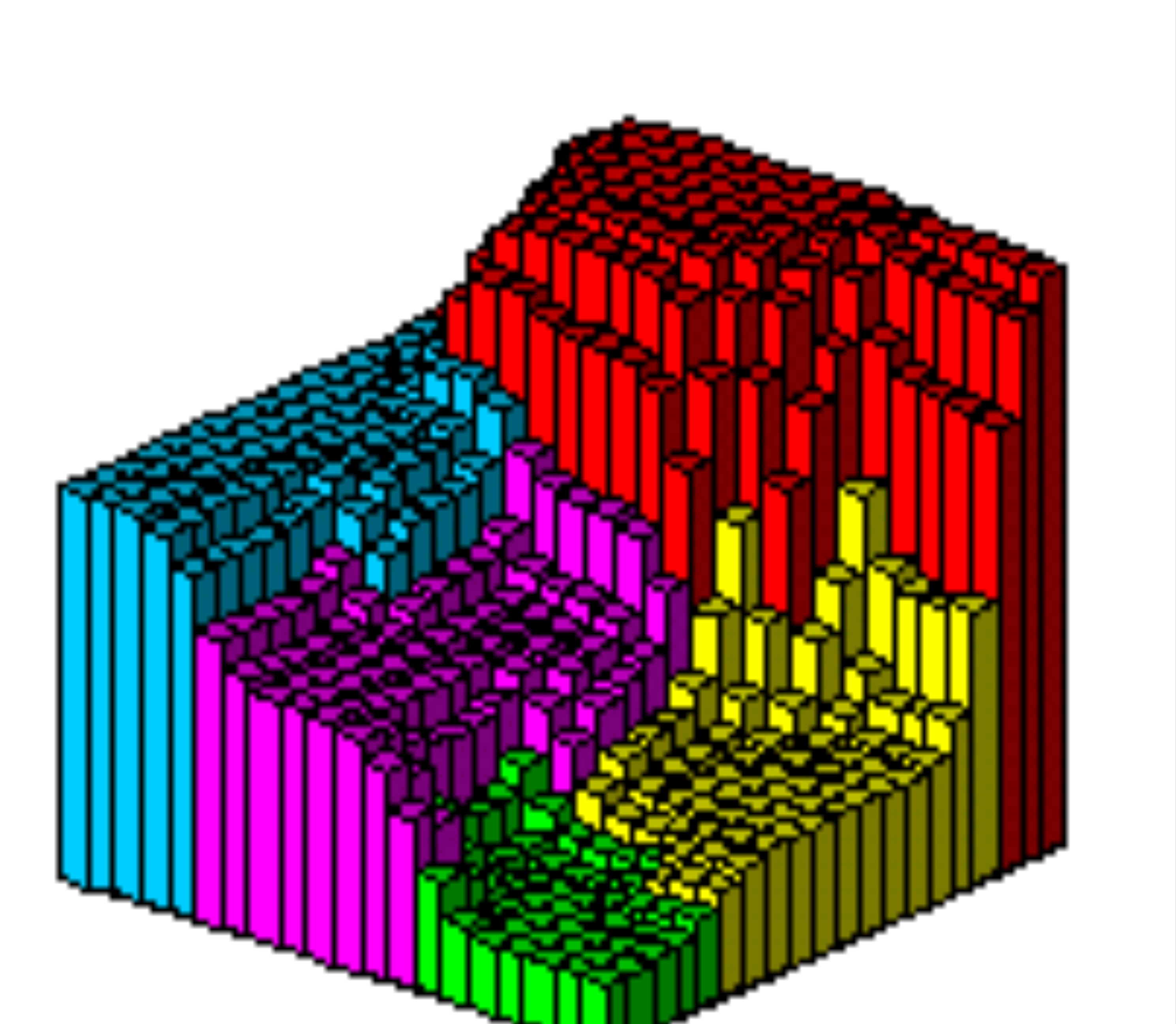
Rasterized polygons of different content



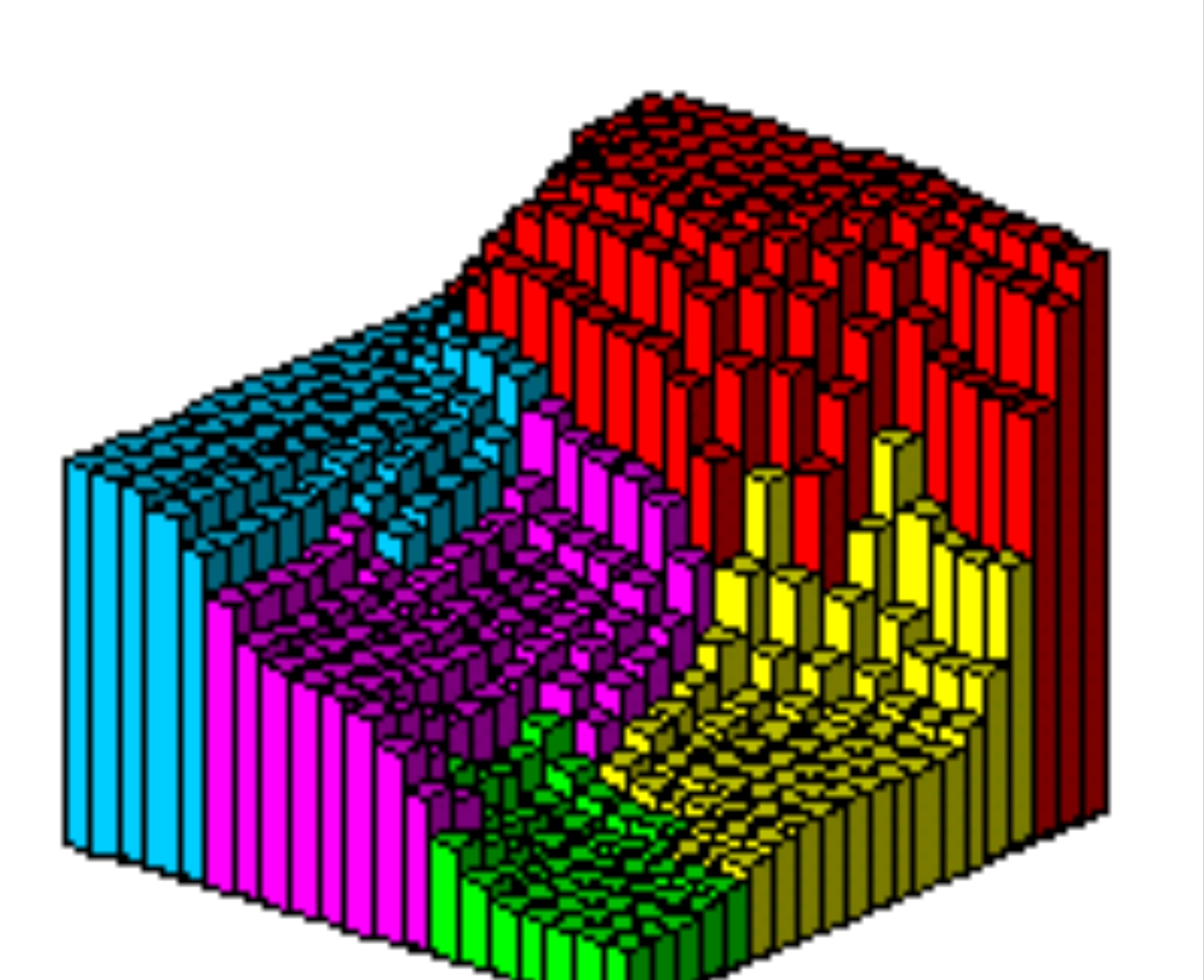
Five Iterations



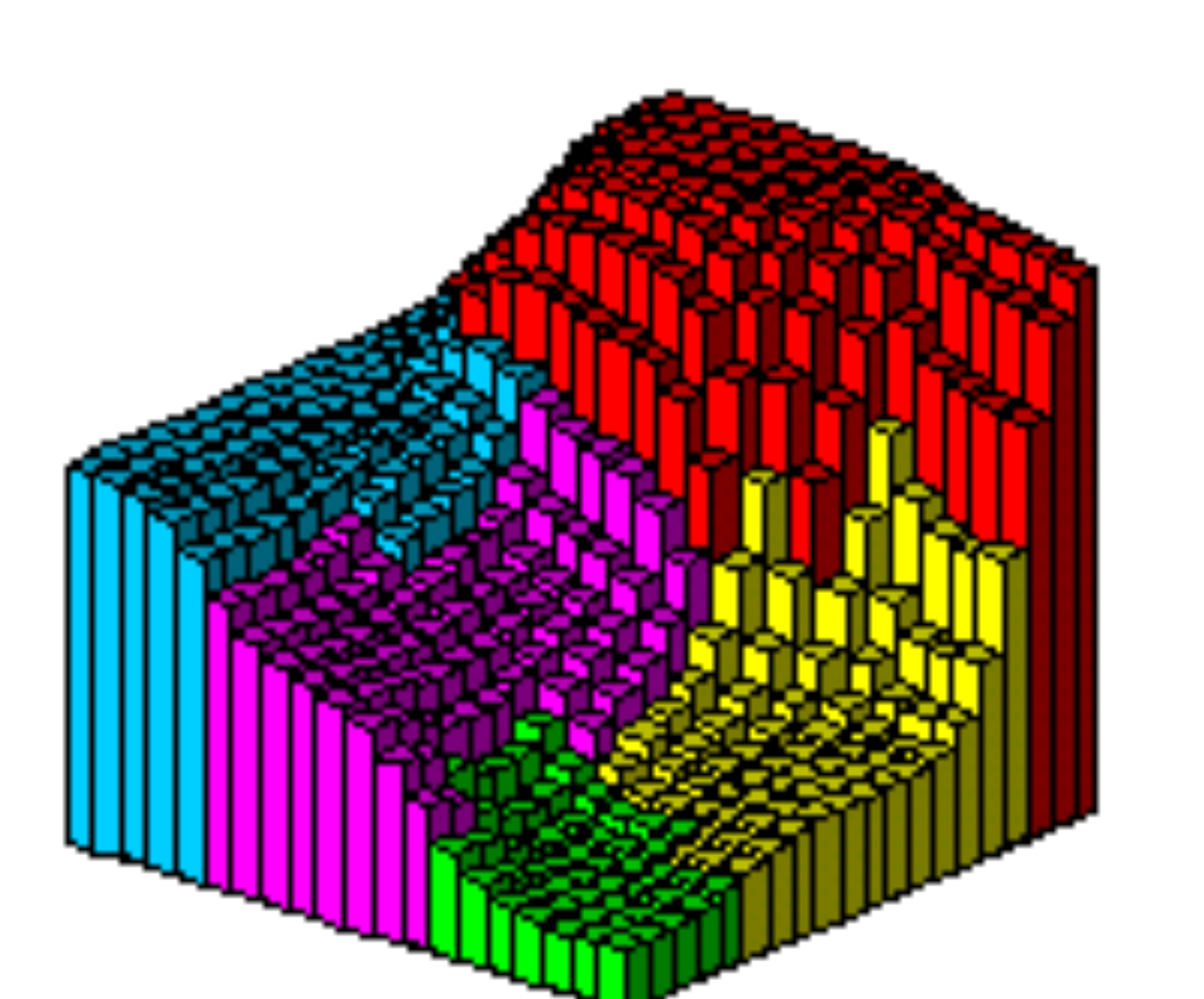
Ten Iterations



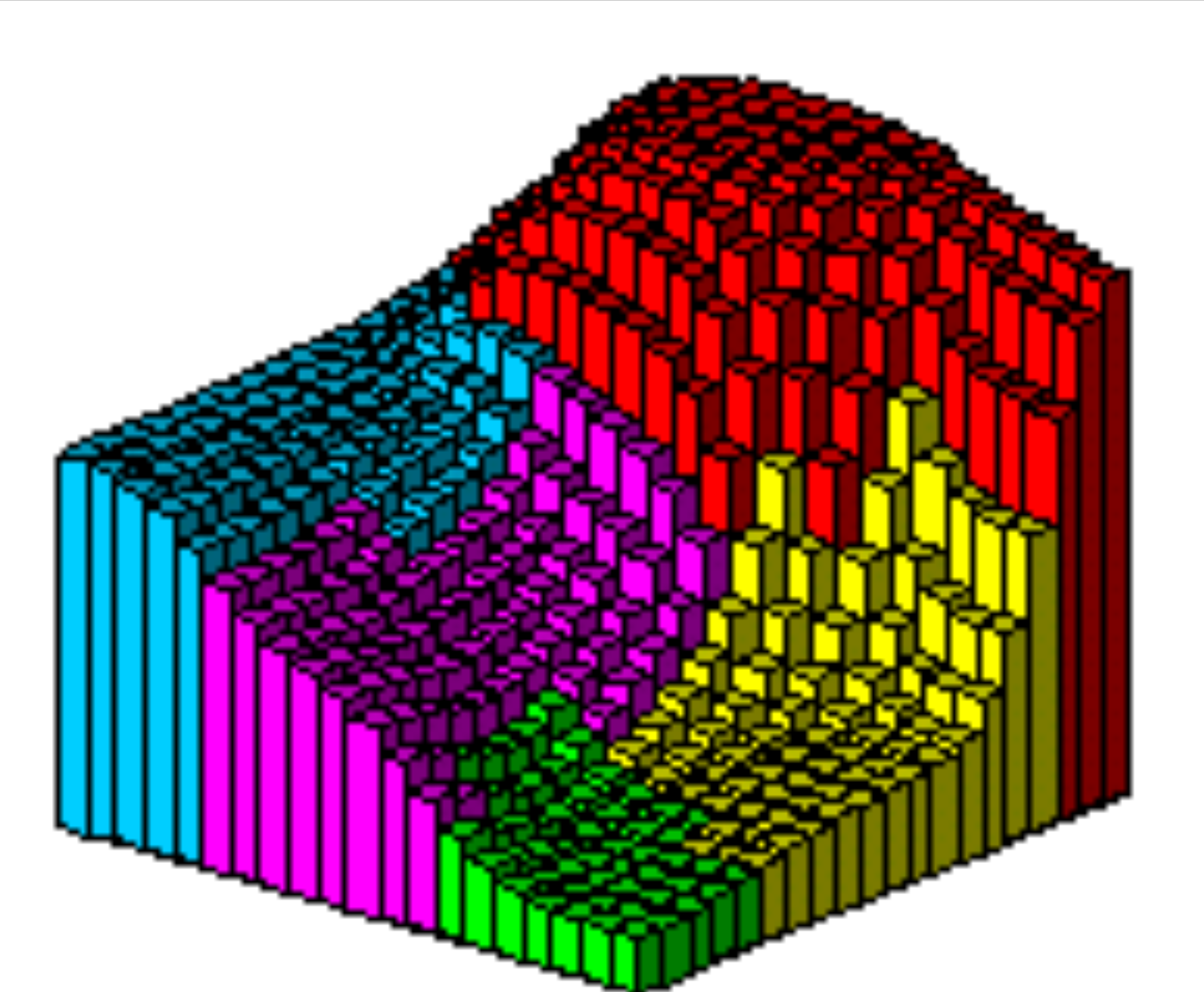
Fifteen Iterations



Twenty Iterations

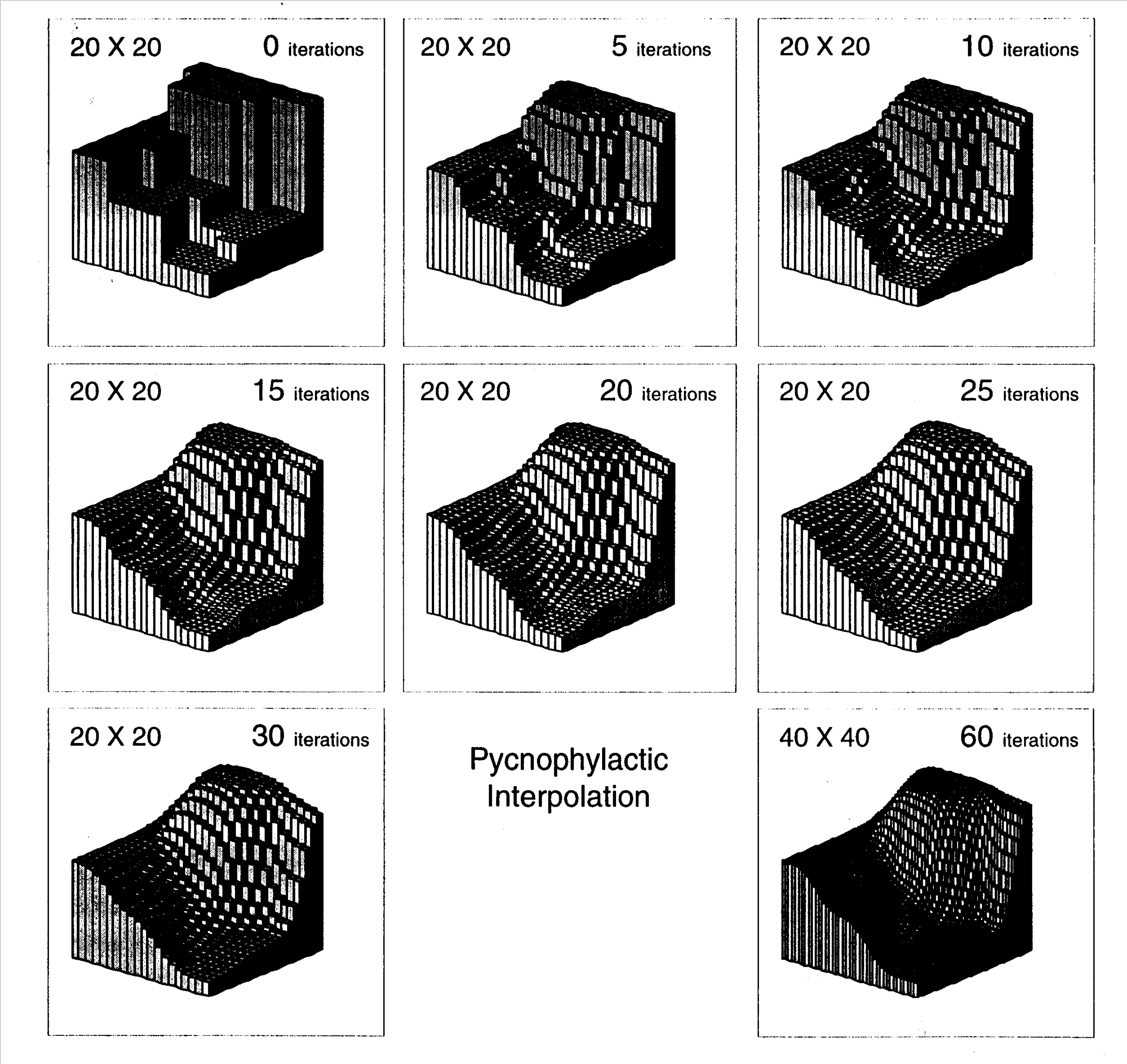


Twenty Five Iterations

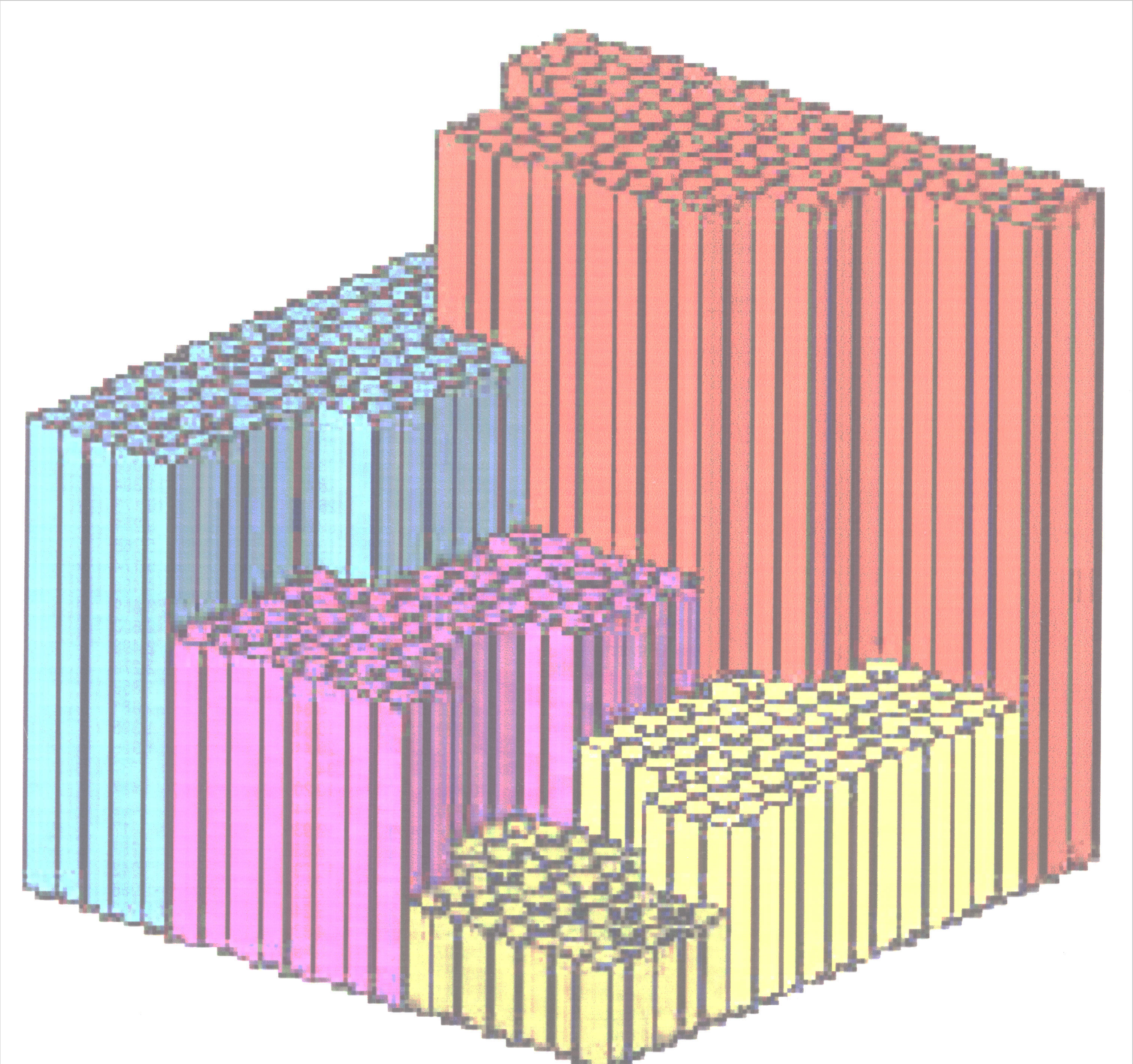


From zero to 60 iterations

The final result is quite smooth

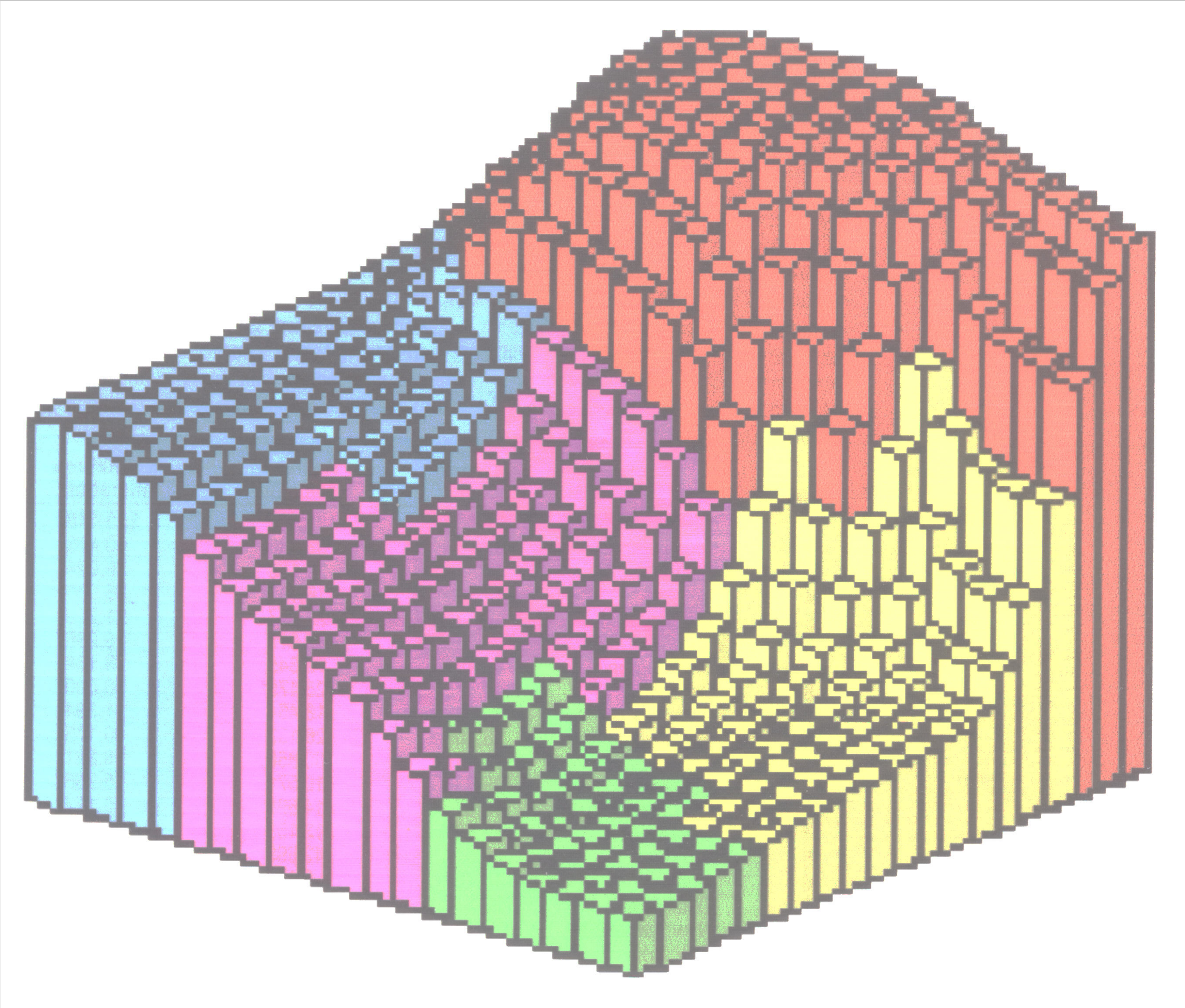


Colored Clay Before Smoothing



Colored Clay After Smoothing

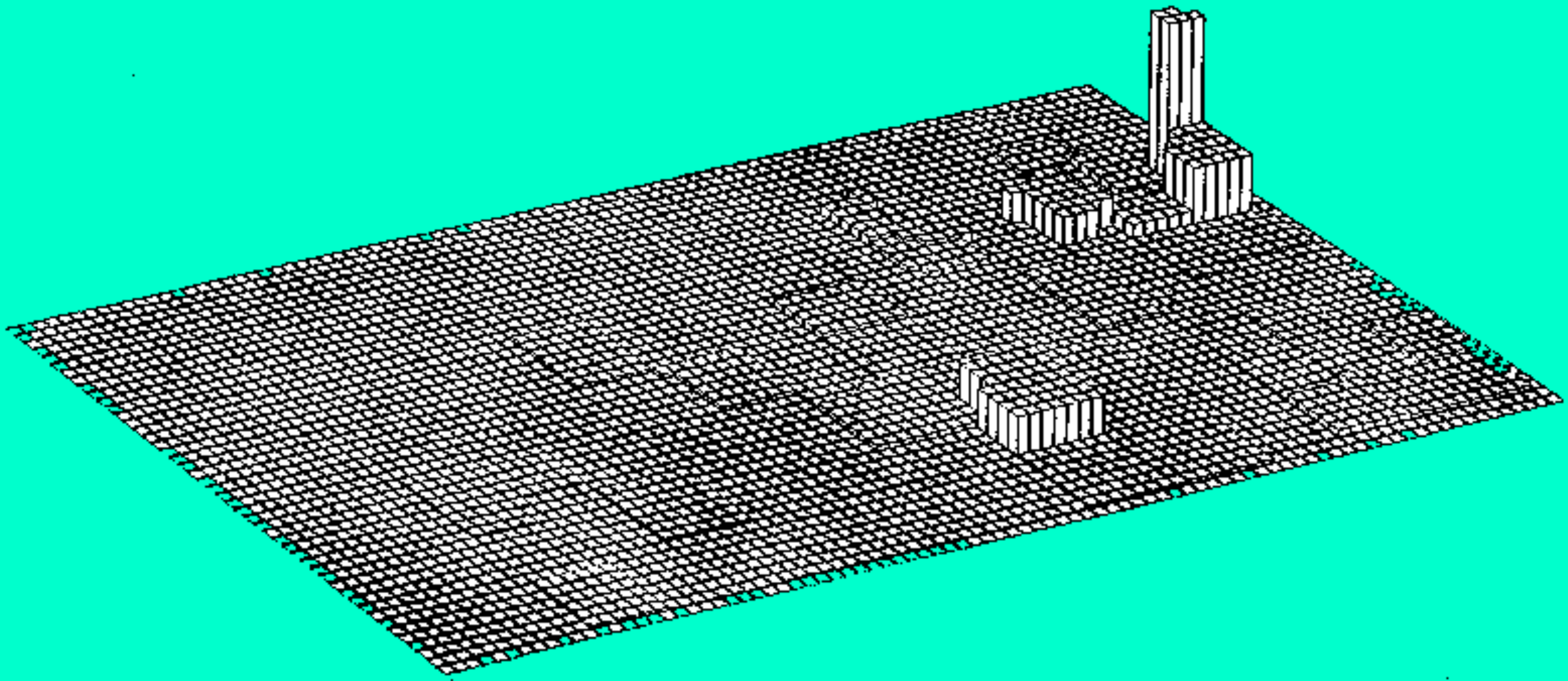
Convert to an alternate set of zones by imposing new boundaries (as if using a cookie cutter), then add up the values in the new polygons



Population Density in Kansas

By County

Courtesy of T. Slocum

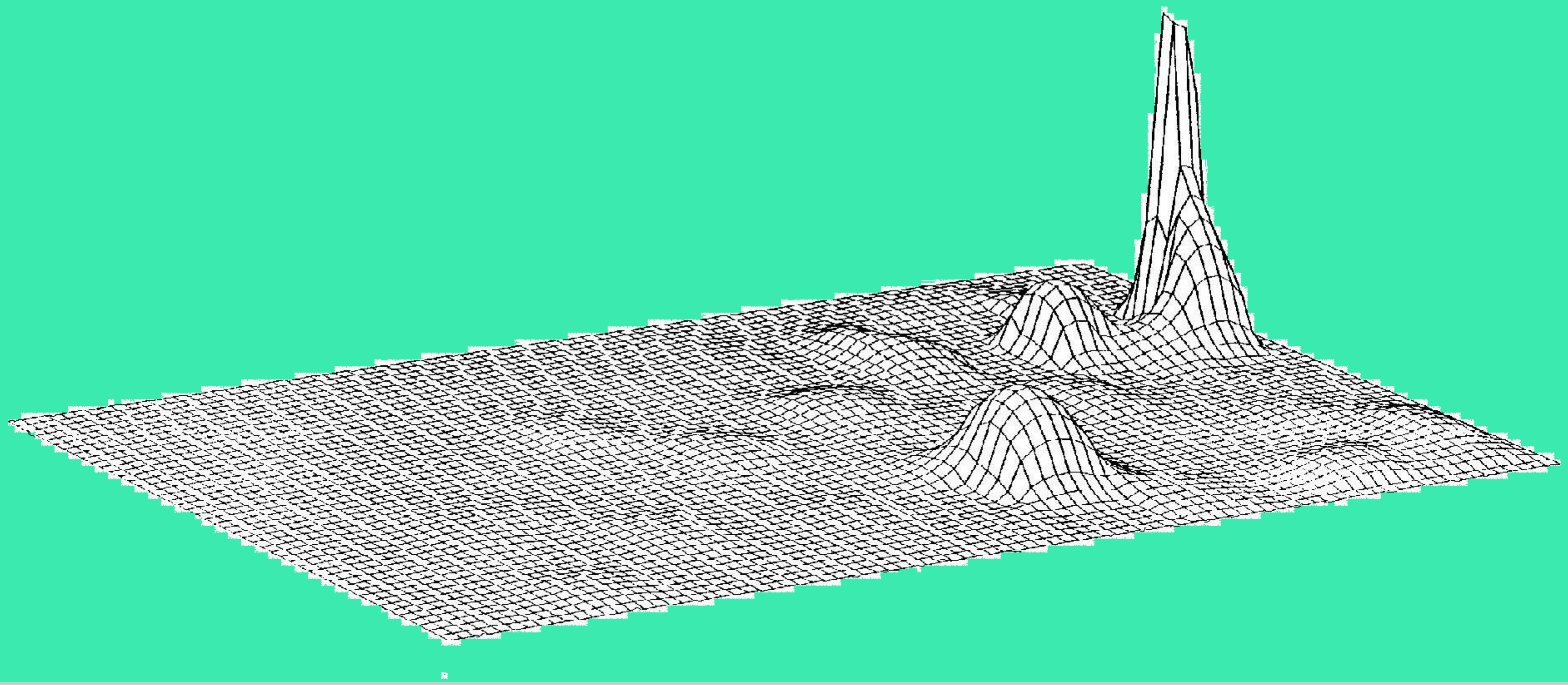


A piecewise continuous surface

Population Density in Kansas

by County

Each county still contains the same number of people



A smooth continuous surface, with population pycnophylactically redistributed

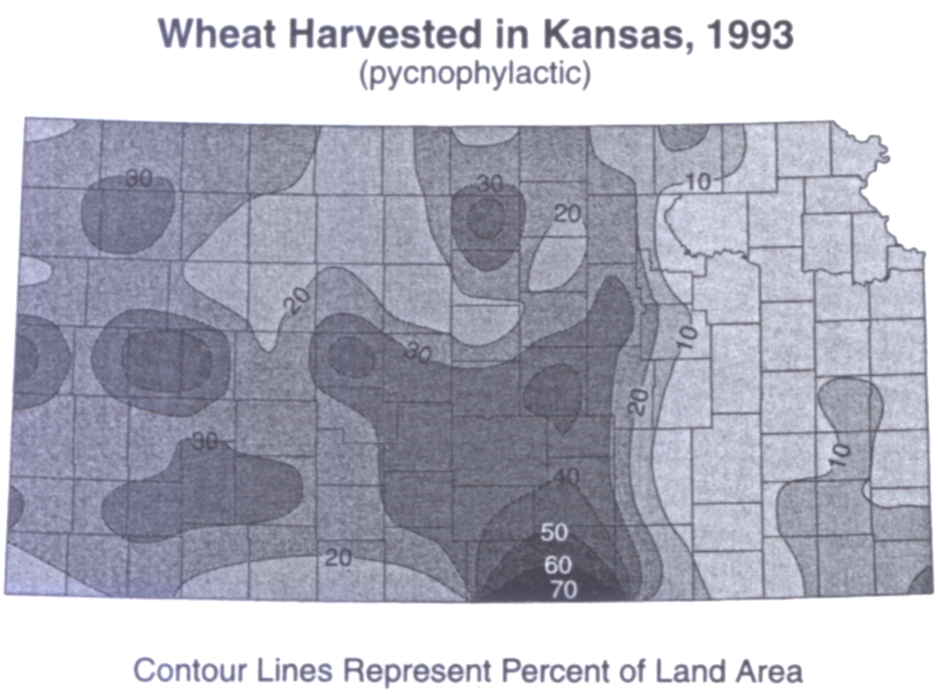
Two types of isopleth interpolation compared.

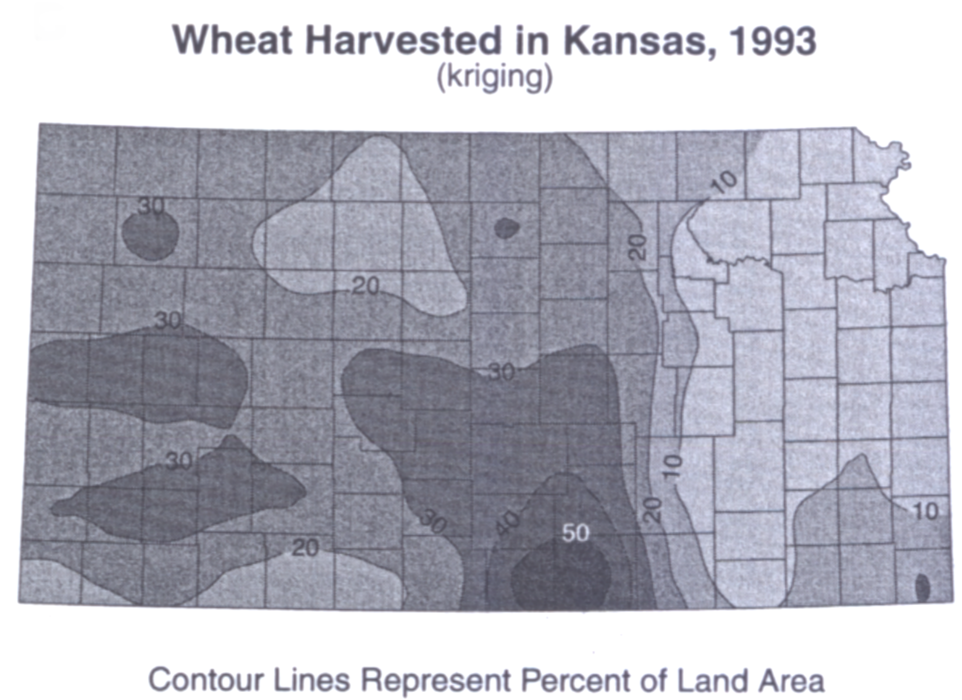
Using data from Kansas counties.

Pycnophylactic reallocation and punctual Kriging from centroids.

The pycnophylactic method does not use centroids, and preserves the county totals.

Figure 8.13 (p. 150) of T. Slocum, "Thematic Cartography and Visualization", Prentice Hall, 1999.





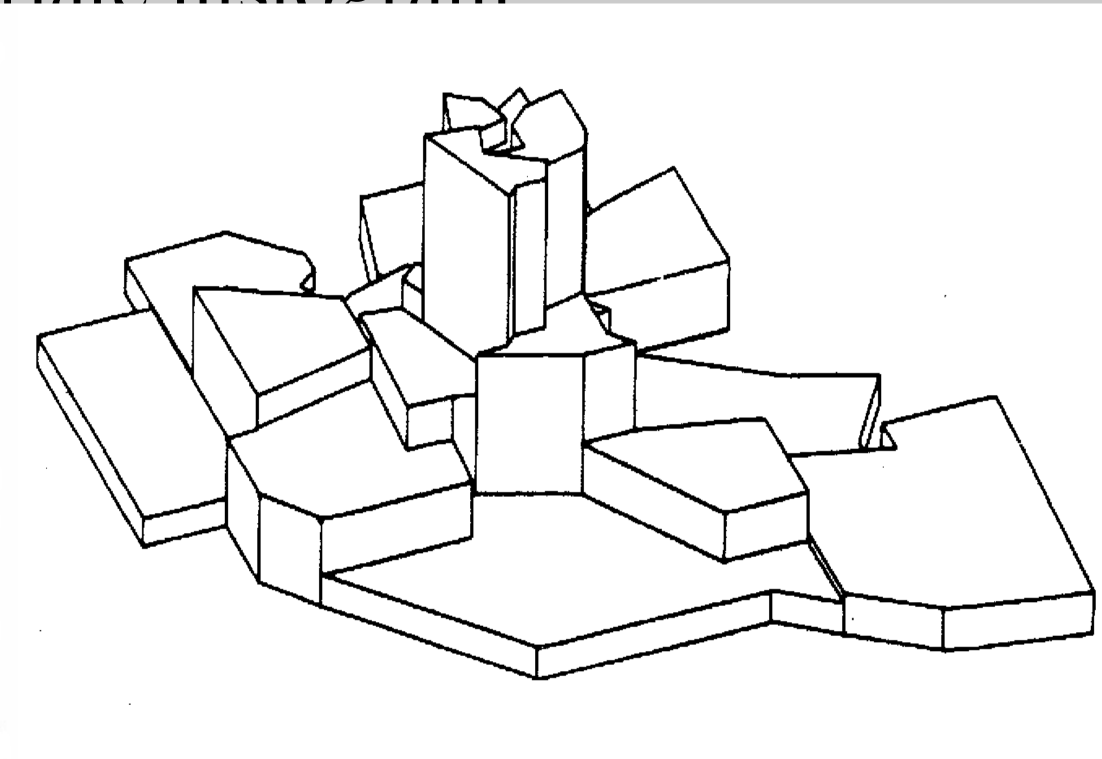
Another Example

Ann Arbor population density by census tract:

Choropleth map

&

Bivariate histogram

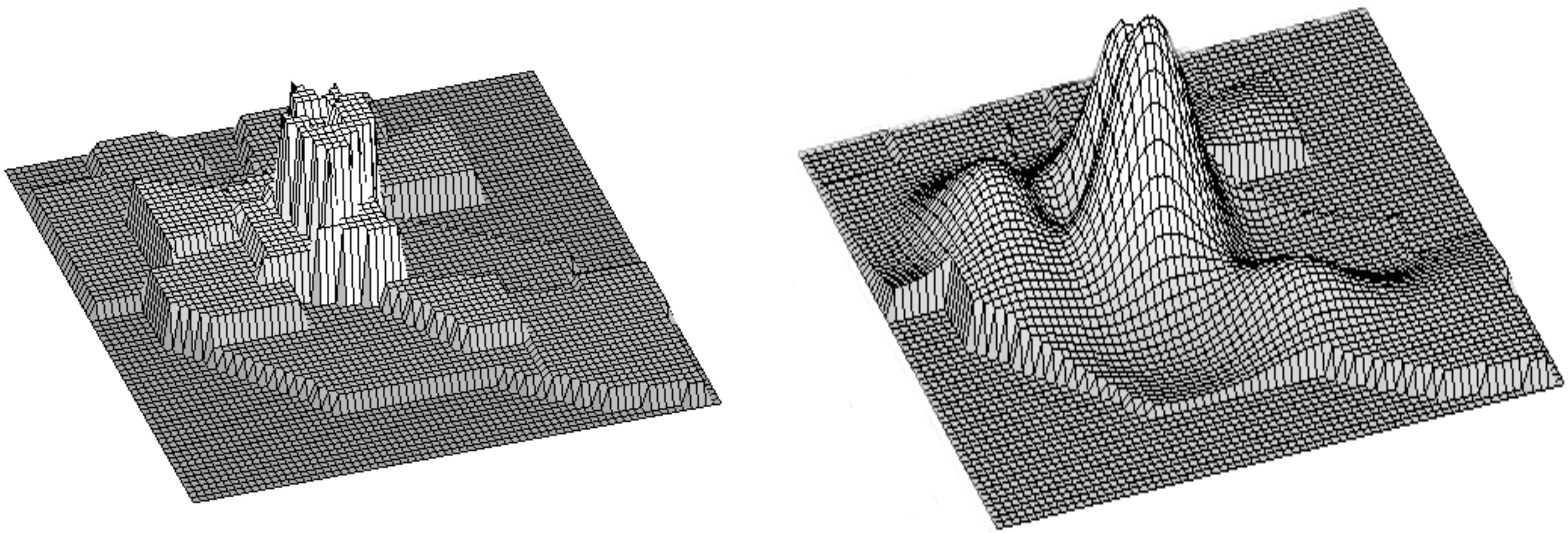


Population density rasterized & reallocated pycnophylactically

zero iterations

&

two hundred iterations

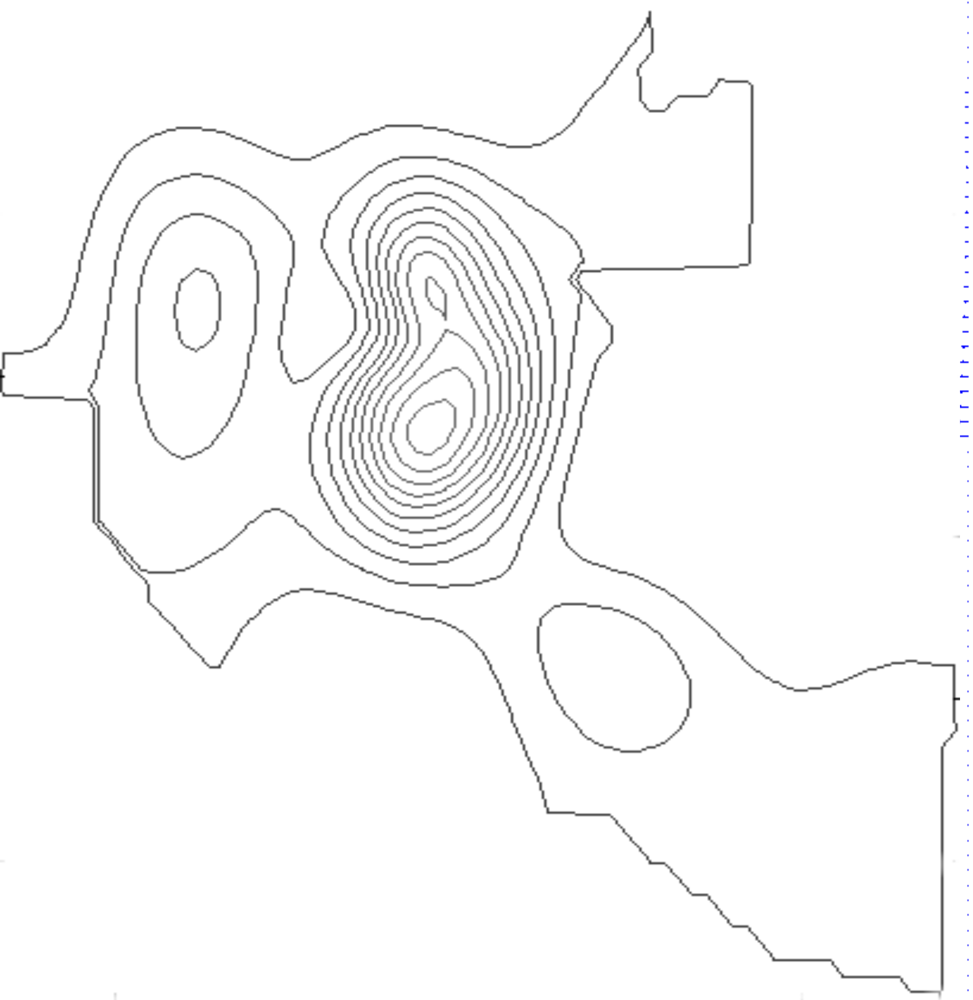


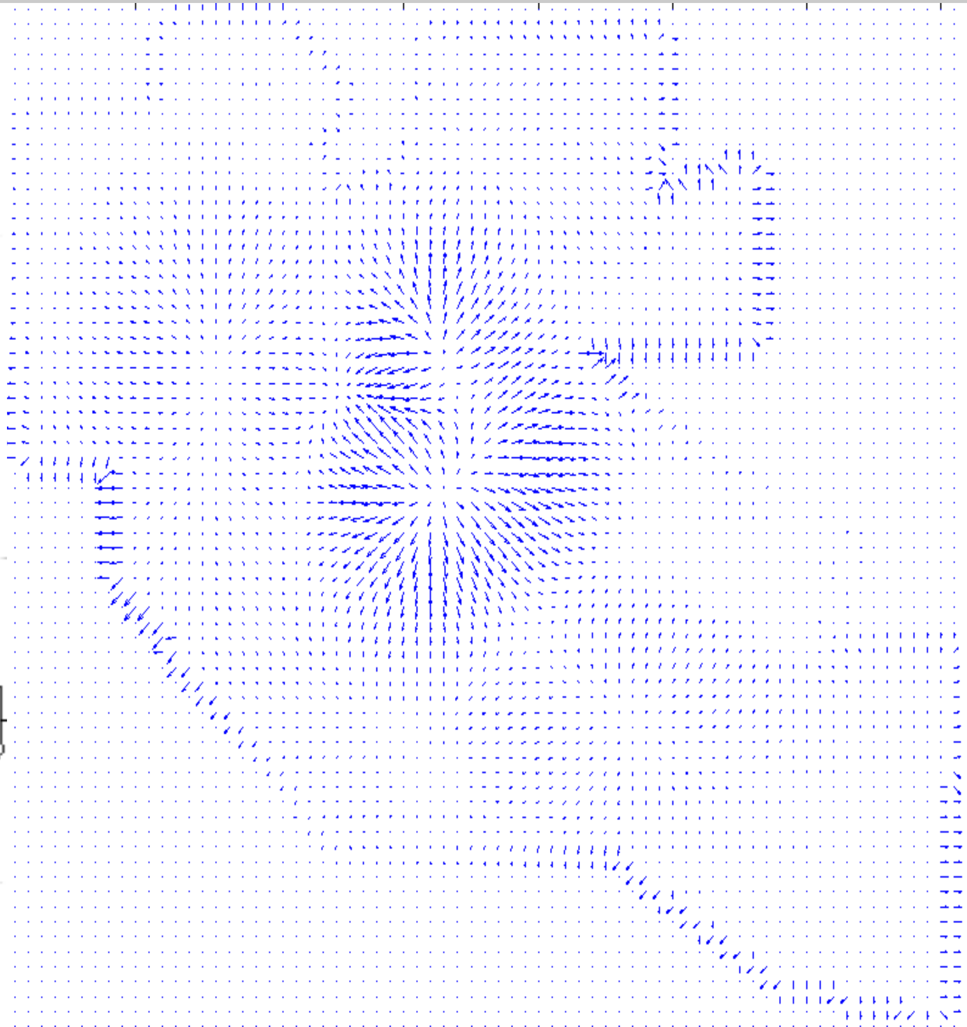
Pycnopylactic reallocation yields a smooth field from which one can make a

Contour map

&

Gradient field



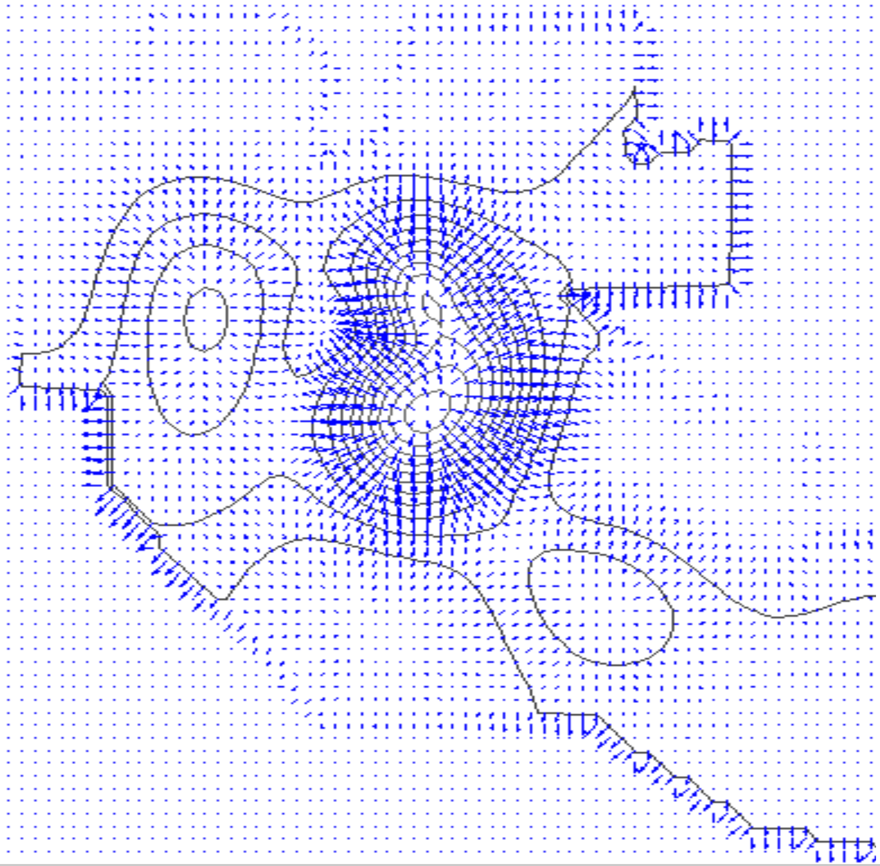


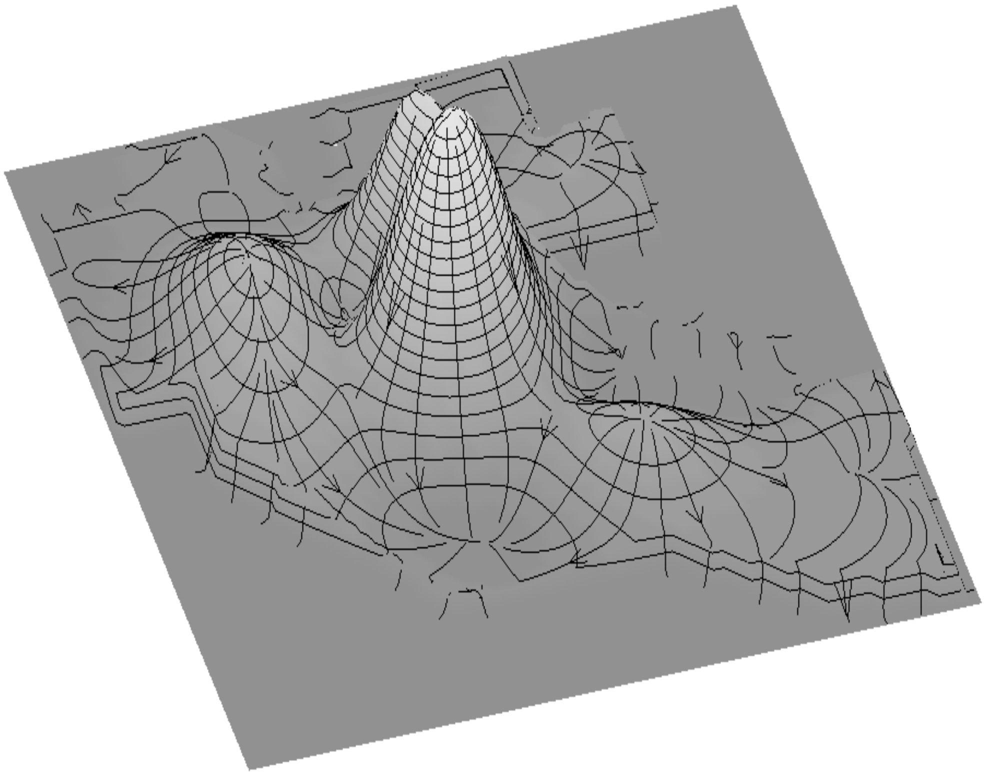
Contours with gradient field

and

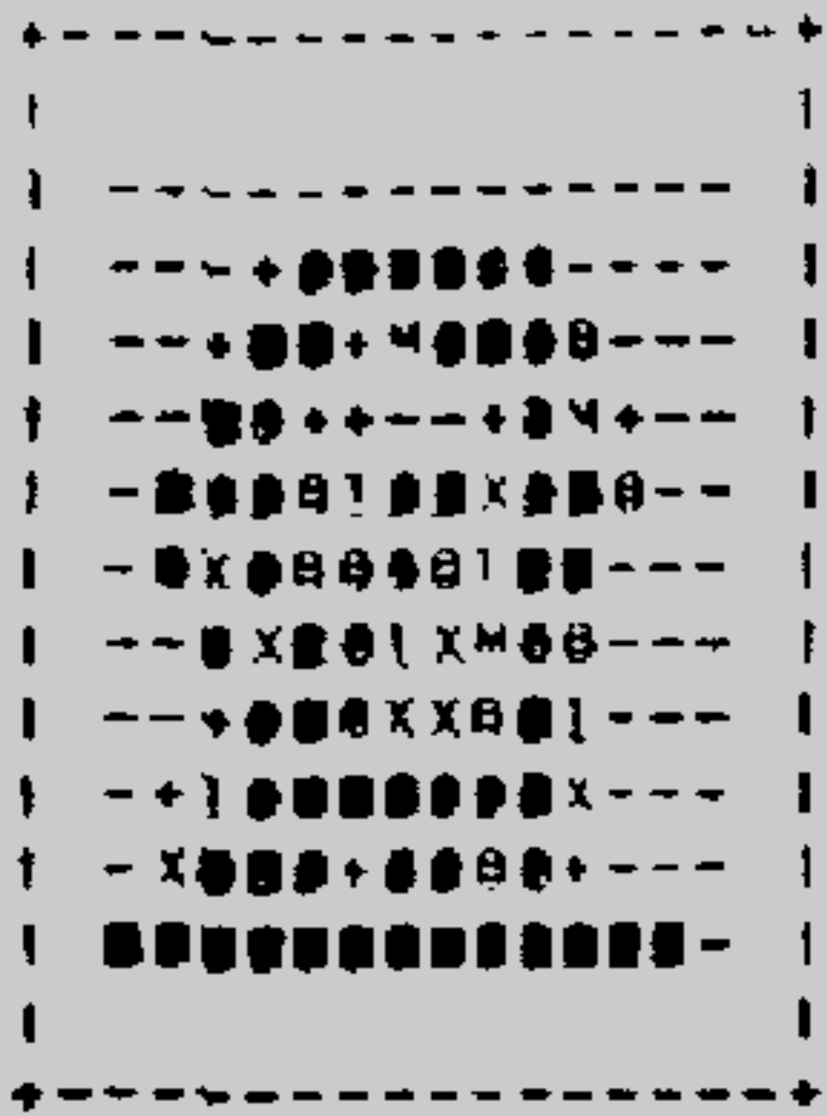
Orthogonals to the contours

Lines of expected growth? (Borchert 1961)





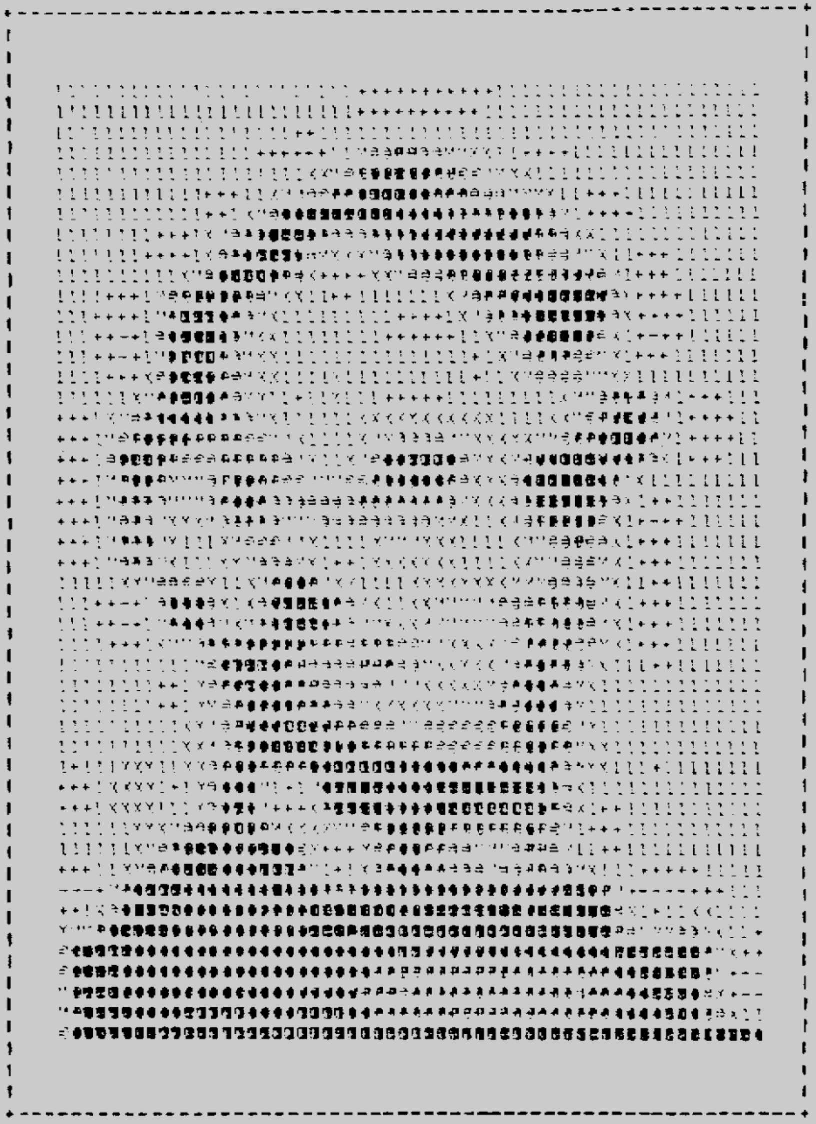
An Image Processing Example
A 20 by 14 Image



Quadrupled from 20 x 14 to 80 x 56

but with the same total "mass"

within each 4 by 4 region



Eleventh Visualization

Computing Potentials from Asymmetry

Many types of data come in the form of square tables.

Often these tables are not symmetric.

Movement, interaction, and communication tables are geographic examples.

Of course there are many other such situations.

It is useful to take advantage of the asymmetry.

An example of a non-symmetric table.

Information movement (citations) among psychology journals

Coombs et al 1970

Data from 1964

	AJP	*JASP*	*JAP*	*JCPP*	*JCP*	*JEdP*	*JExP*	*Pka*	*Total*
American Journal of Psychology	119	8	4	21	0	1	85	2	240
Journal of Abnormal and Social Psychology	32	510	16	11	73	9	119	4	774
Journal of Applied Psychology	2	8	84	1	7	8	16	10	136
Journal of Comparative and Physiological Psychology	35	8	0	533	0	1	126	1	704
Journal of Consulting Psychology	6	116	11	1	225	7	12	7	385
Journal of Educational Psychology	4	9	7	0	3	52	27	5	107
Journal of Experimental Psychology	125	19	6	70	0	0	586	15	821
Psychometrika	2	5	5	0	13	2	13	58	98
Total	325	683	133	637	321	80	984	102	3,265

Let M_{ij} represent a movement table with i rows and j columns. It can be separated into two parts, as follows.

$$M_{ij} = M^{+} + M^{-}$$

where

$$M^{+} = (M_{ij} + M_{ji})/2 \qquad \text{symmetric}$$

$$M^{-} = (M_{ij} - M_{ji})/2 \quad \text{skew symmetric}$$

The percent of variance can be computed for each component,

and the degree of asymmetry can be computed.

How the two parts are used

I consider the symmetric component as a type of background.
The real interest is in the asymmetric part.

In geographic cases the locations are usually known.
In the journal case the position of the places is not known.

Since locations are not given the symmetric part may be used to make an estimate of these positions.

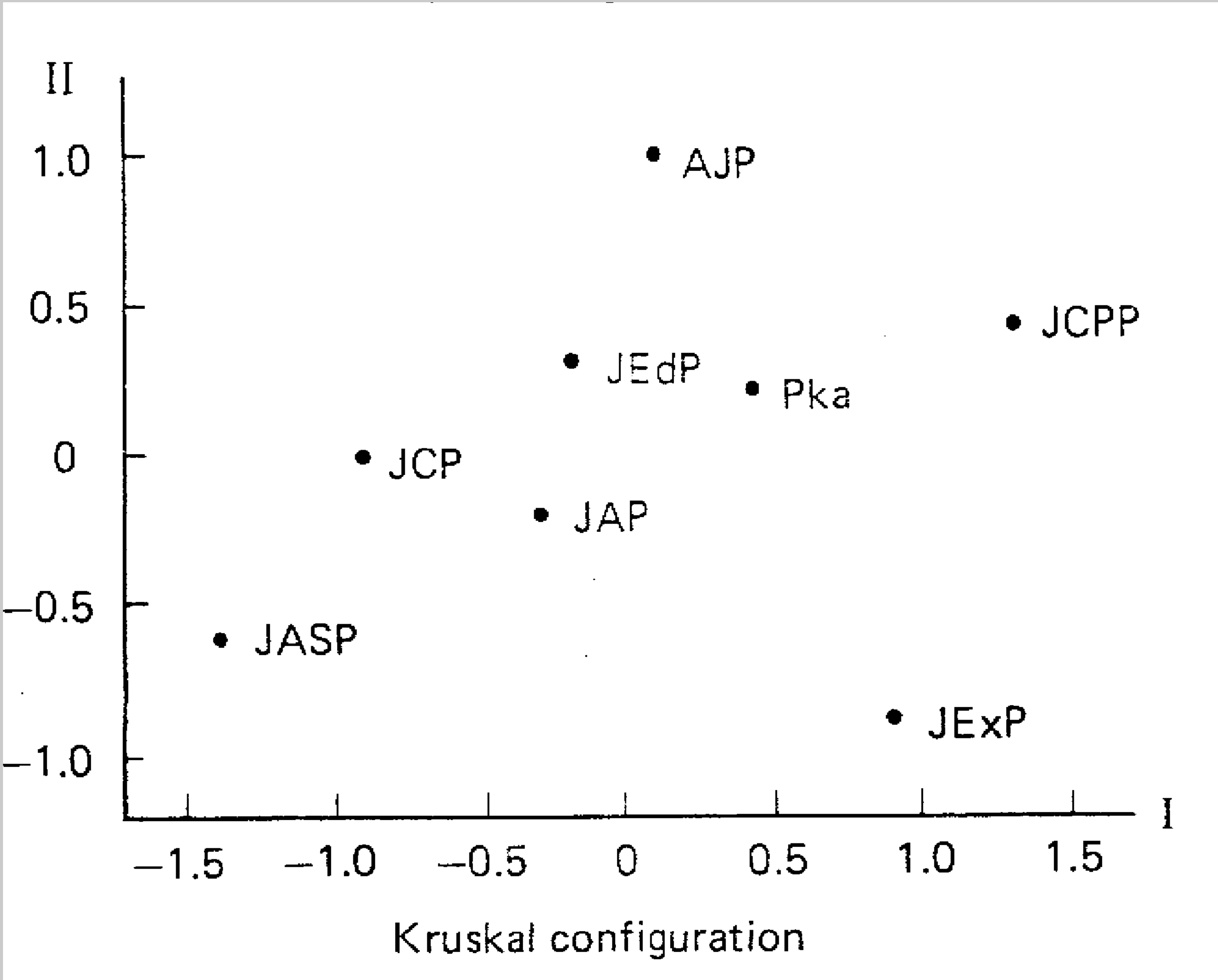
This estimate is made using an ordination, trilateration, or multidimensional scaling algorithm.

The attempt is now made to apply these ideas in a social space.

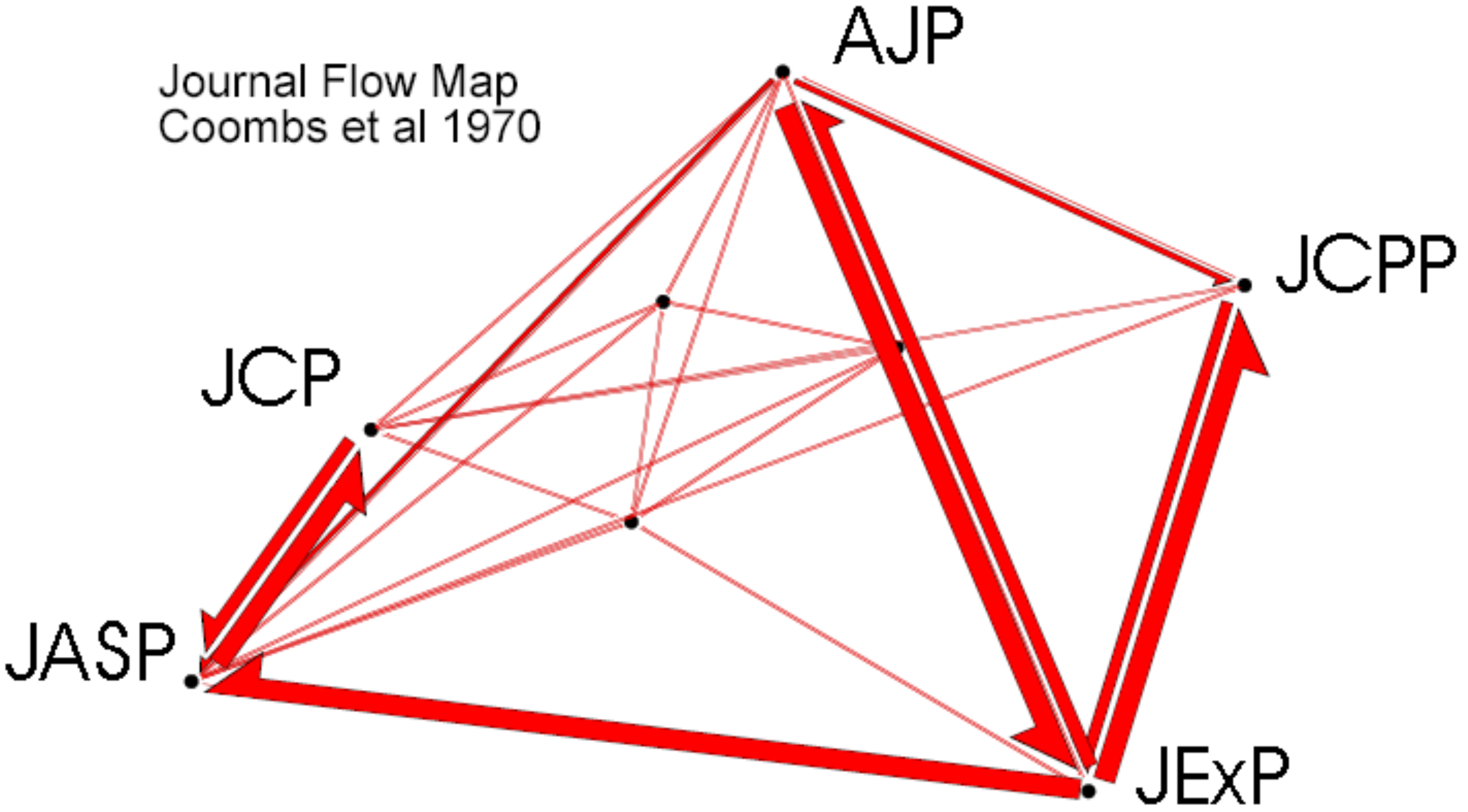
This can be considered a development of Lewin's *Topological Psychology* or his *Field Theory in the Social Sciences*.

The data represent citations between a small set of psychological journals. Larger citation tables are now also available.

In Journal Space



The observed two-way flow between the journals



Journal to Journal Citations

To / From									X	Y	Net Flow
AJP	119	8	4	21	0	1	85	2	125	910	-85
JASP	32	510	16	11	73	9	19	4	-1382	-644	91
JAP	2	8	84	1	7	8	16	10	-261	-237	3
JCPP	35	8	0	533	0	1	126	1	1302	366	67
JCP	6	116	11	1	225	7	12	7	-924	-2	64
JEdP	4	9	7	0	3	52	27	5	-180	324	27
JExP	125	19	6	70	0	0	586	15	904	-924	-163
Pka	2	5	5	0	13	2	13	58	416	207	-4

AJP Am J of Psychology

JASP J of Abnormal & Social Psychology

JAP J of Applied Psychology

JCPP J of Comparative & Physiological Psychology

JCP J of Consulting Psychology

JEdP J of Educational Psychology

JexP J of Experimental Psychology

Pka Pyschometrika

C. Coombs, J. Dawes, A Twersky, 1970, *Mathematical Psychology*, Prentice Hall, Engelwood Cliffs, NY, Pages 73-75

The table gives the being-cited journal across the columns. But the information can be considered to move from that journal to the citing journal.

Therefore the transpose is used to produce the source to sink map.

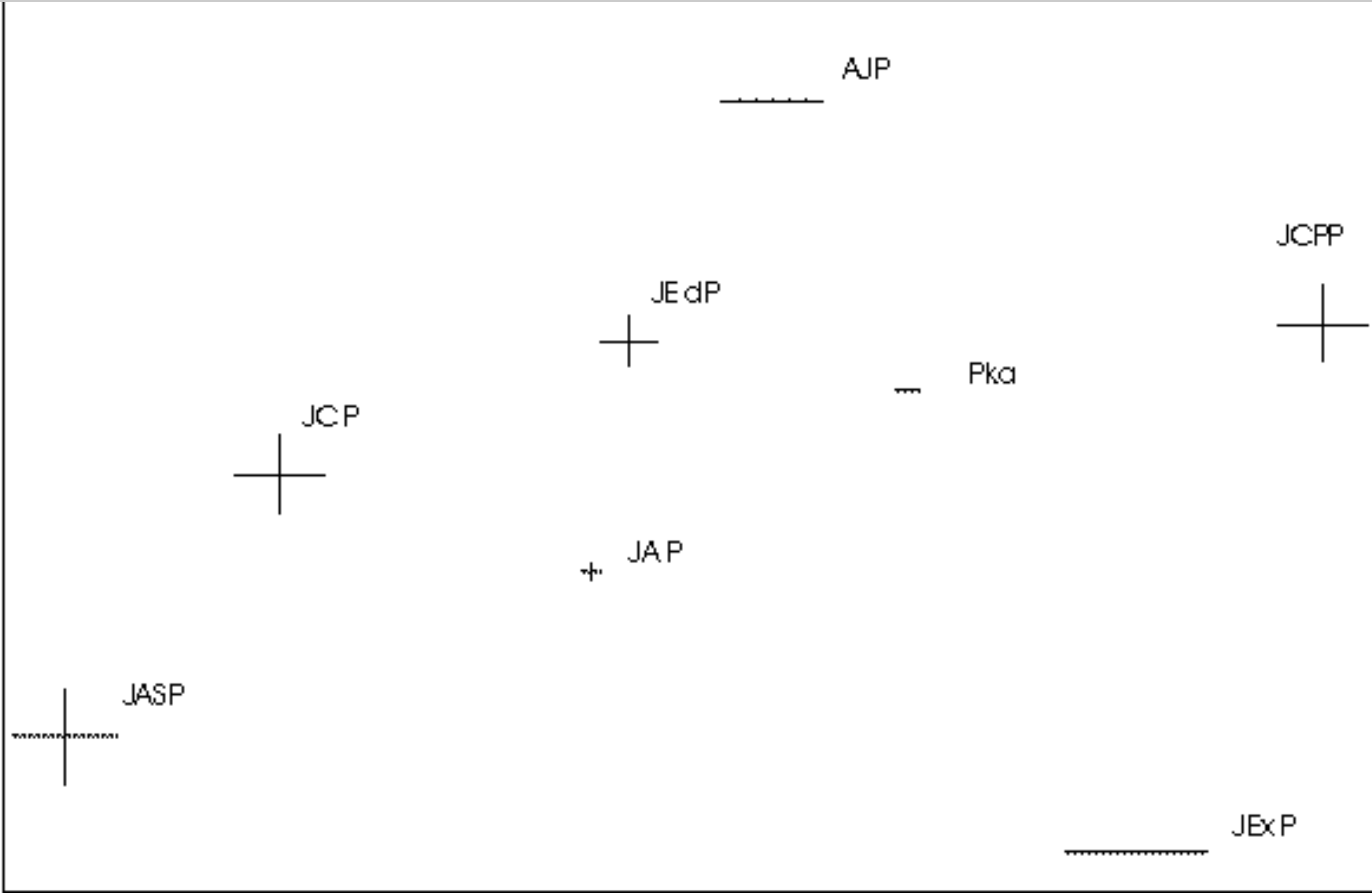
Adding across the table, the column marginals give the outsums (a.k.a. outdegree). Summing down the rows gives the insums (a.k.a indegree).

The 'sending' places (rows) are known as 'sources', and are shown by negative signs.

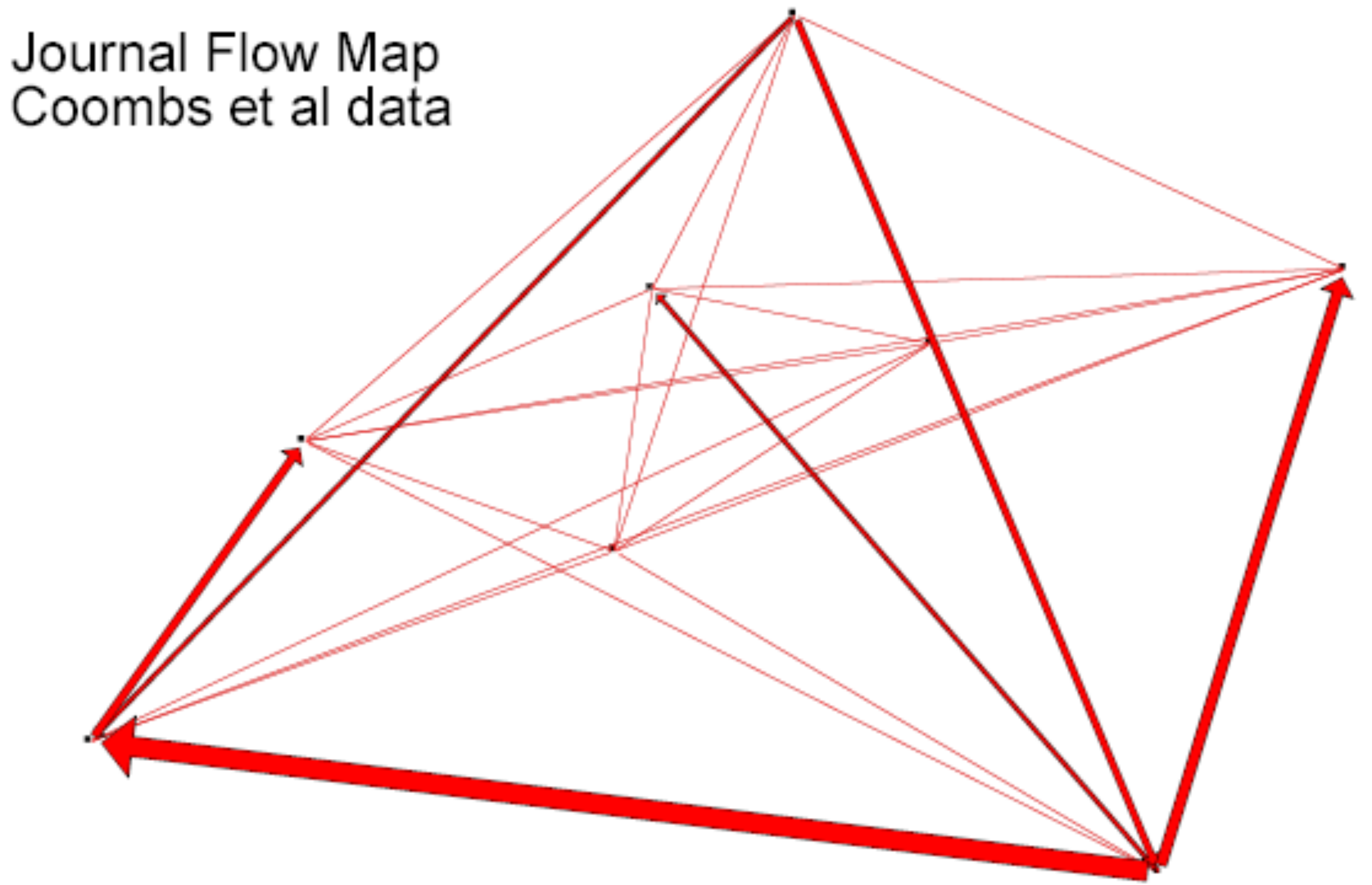
The 'receiving' places (columns) are the 'sinks', and are shown by plus signs.

The size of the symbol represents the magnitude of the movement volume.

Journal Sources and Sinks



The net flow between the journals



We now have an assignment problem. How to get 163 citations from JExP, 85 from AJP, & 4 from Pka to the 5 receiving journals, using only the marginals. There are obviously many possibilities

One solution is to use the "Transportation Problem" (Koopmans, Kantorovich, ~1949): Minimize M..d.., subject to $M_{.J} = O_I$, $M_{I.} = I_J$, $M_{IJ} >= 0$, given the distances computed from the coordinates and using the simplex method for the solution.

A more realistic solution is given by the quadratic transportation problem: Minimize M^2..d.., subject to the same constraints.

Both of these solutions result in discrete answers, and 'shadow prices'. I am looking for a spatially continuous solution that allows vectors and streamlines, in order to determine spatial flow fields and a continuous potential.

The next step is to compute the displacements between the cited journals.

This is based on the asymmetry of the citations table.

The fundamental idea being that there exists a 'wind' making movement easier in some directions.

The mathematical details are given in a published paper.

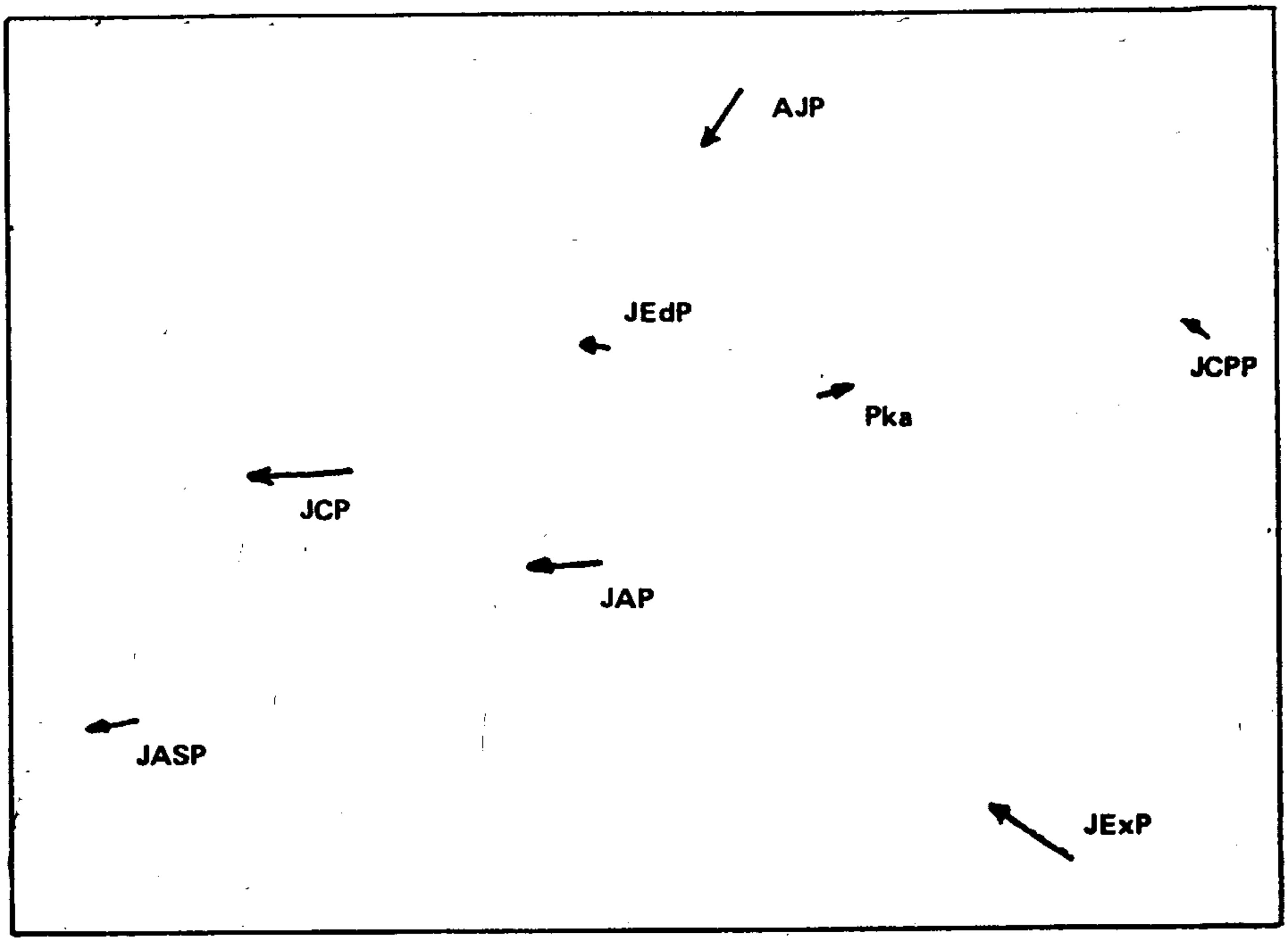
W. Tobler, 1976, "Spatial Interaction Patterns", *J. of Environmental Systems*, VI(4):271-301

The movement from source to sink can be computed to show the direction and magnitude of the movement.

The computation is based on the asymmetry of the movement table.

Small directed vectors represent this movement on the next map.

Displacement between Journal Citations



An interpolation is then performed to obtain a vector field from the isolated individual vectors.

This is done to simplify the mathematical integration needed to obtain the forcing function.

The computed potential should have the vector field as its gradient.

This is a hypothesis that can be tested.

The base level of the potential is determined only up to a constant of integration.

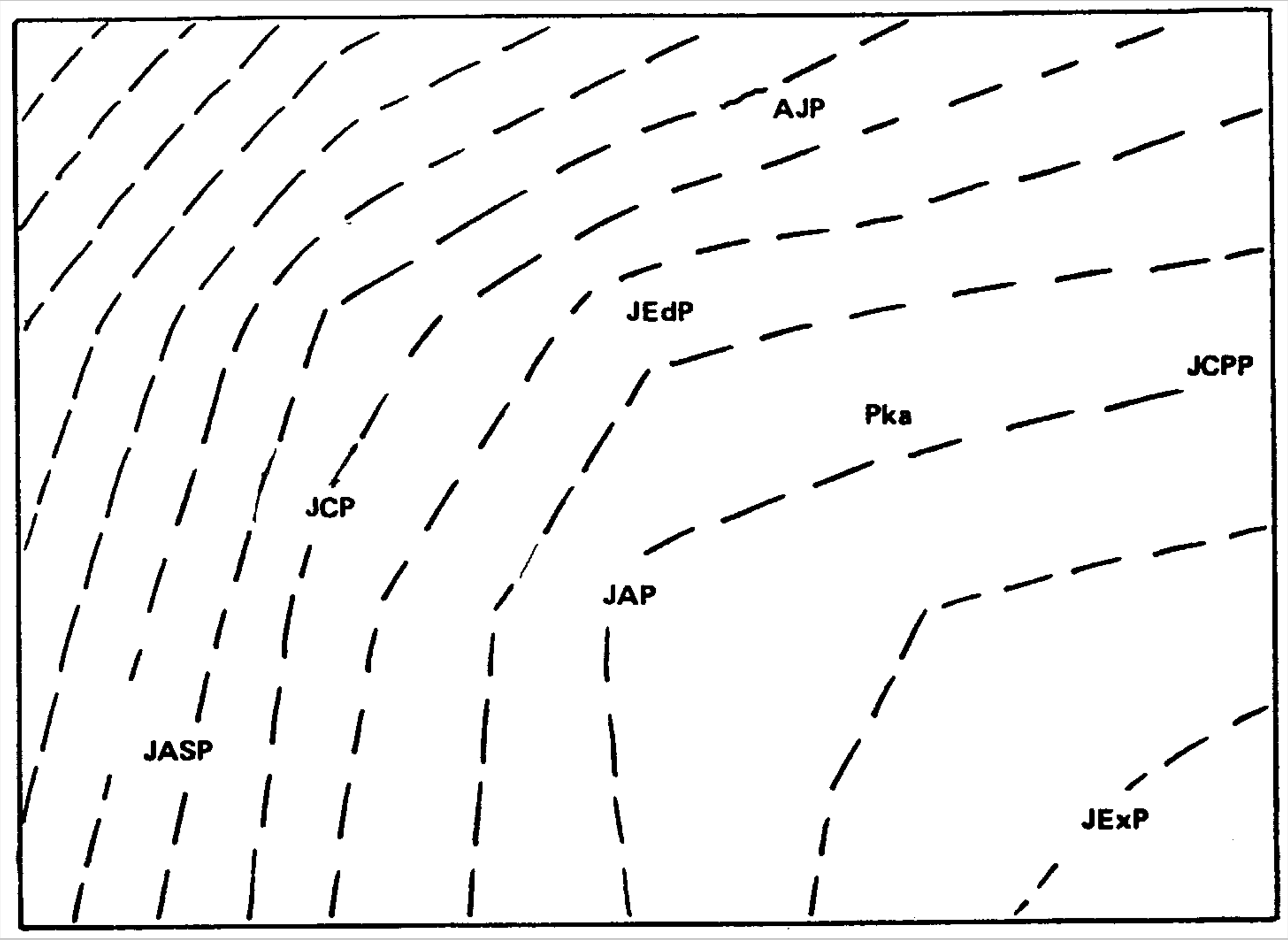
The vector field, to be a gradient field, must be curl free. This can also be tested.

Then the potential is computed by integration.

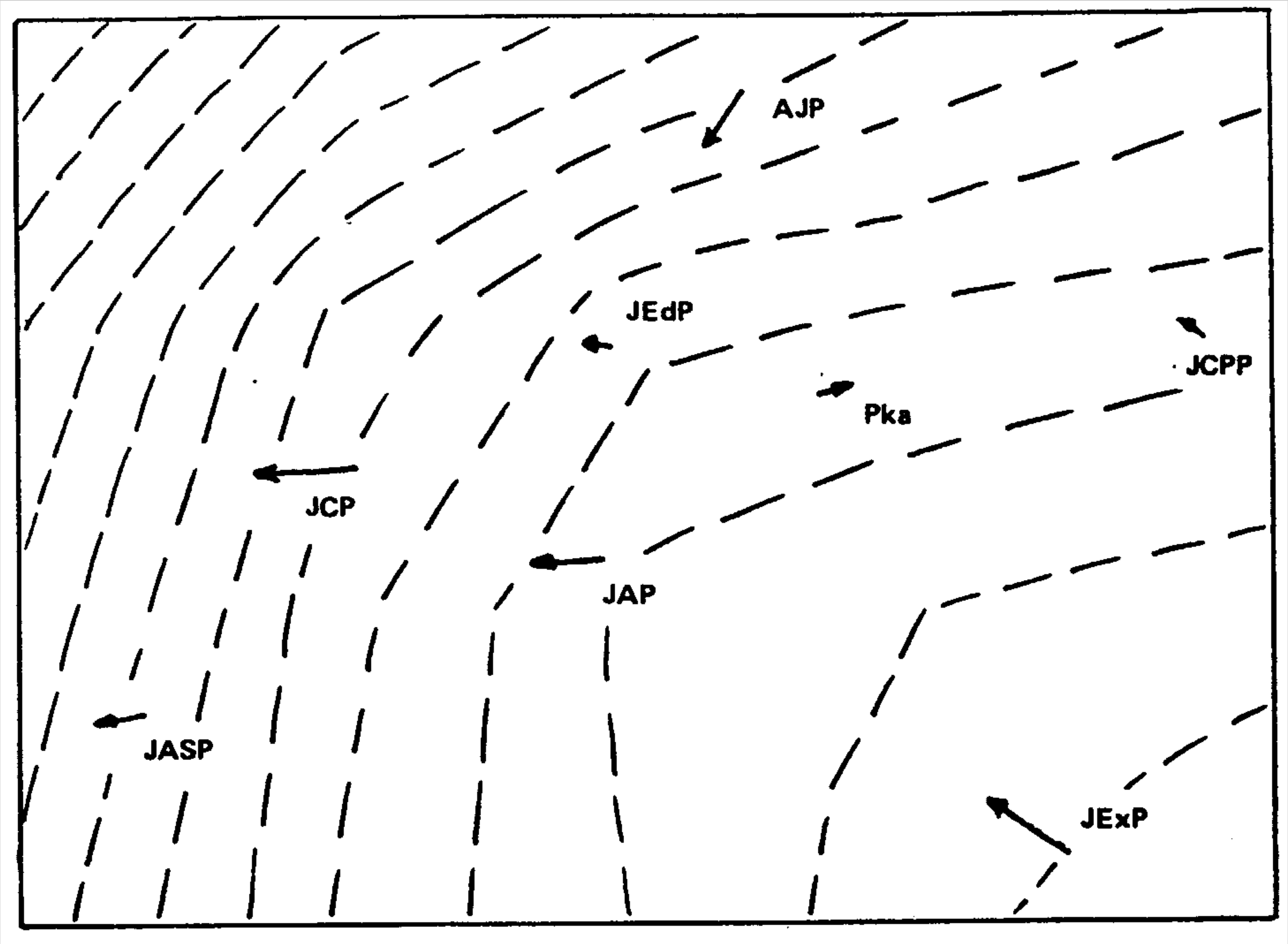
This potential should be such that its gradient coincides with the displacement vectors.

It may be necessary to use an iteration to obtain this result.

Journal Potential Function



Flow and Potential between Psychological Journals



Some questions

Suppose a new psychological journal were started. Where should it be inserted into in this space?

Does it make sense to treat journal citations as being located in a continuous two-dimensional social space?

Can other social data be treated in a similar fashion, for example social mobility tables?

And more general network data?

Conclusion

Spatial studies are readily adaptable to graphic visualization.

This is because they involve geometry.

Graphics often clarify concepts that are otherwise difficult to understand

and dramatically suggest topics for further study.

This is well known but often under-utilized by analytical scientists.

For more examples go to
http://www.geog.ucsb.edu/~tobler

Thank you for your attention

Eleven examples were demonstrated.

1: Spatial and Temporal Autocorrelation.
2: Comparison of movement patterns.
3: Changing Resolution.
4: Migration potential.
5: Biproportional Adjustment.
6: Coalescent cities.
7: Multidimensional Scaling Iterations.
8: Transform-solve-invert
9: Vector Map Matching.
10: Pycnophylactic Reallocation.
11: Potential from Asymmetry

Questions?

Some Readings

H. Funkhouser, 1937, “Historical development of the graphical representation of statistical data”, Osiris, 3:269-404.

M.Friendly, D. Denis, ~2008, Milestones in the history of thematic cartography, statistical graphics, and data visualization. http://datavis.ca/milestone/

T. Hankins, 1999, “Blood, Dirt, and Nomograms”, Isis, 90:50-80.

G. Palsky, 1996, “Des Chiffres et les Cartes”, Comite des Traveaux Historiques et Scientifiques, Paris.

H. Wainer, 2005, Graphic Discovery, Princeton University press.

L. Wilkinson, 1999, The Power of Graphics, New York, Springer.