Methods of modelling in the light of information theory and entropy.

Paul Cockshott

I. Modeling: a materialist approach

A scientific model is a machine for predicting how part of reality will behave. We tend to think of models in an abstract conceptual sense, I want to argue that we should look at them in a very concrete material sense. I will present historical examples, and then formulate criteria for the 'goodness' of models, before applying these general principles to economic modeling. The whole approach is very computational and borrows heavily from information theory.

The basic modeling process : is shown in Figure I.1. We have the world and we have a model. The model is, at one level a 'black box' which gives us answers about the world. Given some initial observations we can feed them into the black box and it comes out with predictions. The predictions are not necessarily predictions in time, they may be predictions about other things in the real world that we have not yet observed. If we make the other observation we can then compare what the model predicted with what has actually occurred. At a global scale our model might be a general climate model predicting the long term path of climate under CO_2 emission forcing. The initial observations could be the trends in CO_2 emissions, and the predictions could be ones about future average temperatures and rainfall.

When building such a model one could take past periods say 1900 to 2000 as the initial observations and have it make predictions about climate for the period 2001 to 2009. Generally we work on the rule that the closer the predictions to the other observations, the better is the model.

Are models 'ideas' or are they machines?: The term model is used in two rather different senses. On the one hand we speak of a model as something physical. A model boat may server as an adult's toy, but for a shipyard it also has a serious predictive value. In a wave tank the model tells us something about the performance of the finished boat. Measurements from models were, in the past, used to guide the construction of actual steel plates. So it is clear that models in this everyday meaning of the word, have a real pragmatic use. On the other hand a model is used to refer to something conceptual. One talks of the Newtonian model of the solar system as opposed to the Ptolemaic one. In this case we think of the model as something abstract, an idea not a thing. And this gives rise to all sorts of questions about how it is that a conceptual or mathematical model can be so good at predicting the real world. Wigner complained about the 'unreasonable effectiveness' of mathematics, Hamming asked how simple maths could be so effective. If we think of things in this way, as a correspondence between two quite different domains – that of thought and that of reality – the whole process seems so remarkable as to tempt one to ascribe some mystical properties to mathematics.

> There is a story about two friends, who were classmates in high school, talking about their jobs. One of them became a statistician and was working on population trends. He showed a reprint to his former classmate. The reprint started, as usual, with the Gaussian distribution and the statistician explained to his former classmate the meaning of the symbols for the actual population, for the average population, and so on. His classmate was a bit incredulous and was not quite sure whether the statistician was pulling his leg. "How can you know that?" was his query. "And what is this symbol here?" "Oh," said the statistician, "this is pi." "What is that?" "The ratio of the circumference of the circle to its diameter." "Well, now you are pushing your joke too far," said the classmate, "surely the population has nothing to do with the circumference of the circle." Wigner (1960)Modelling: a materialist approach

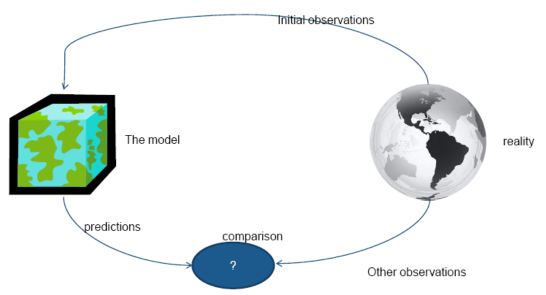

Figure I.1. The basic modeling process

Figure I.2. A model may be understood as a reduced physical copy.

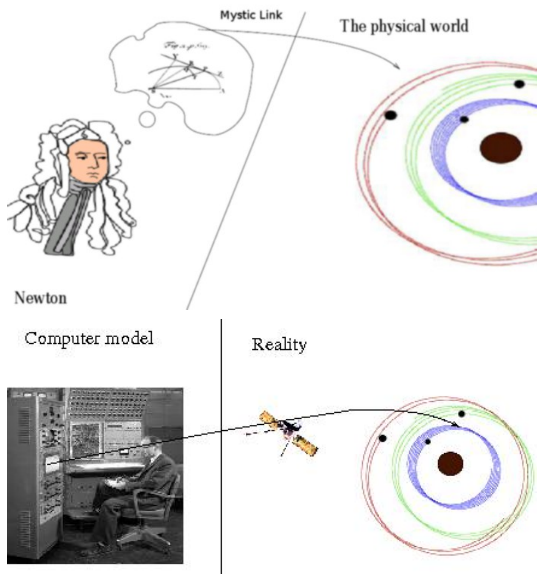

Figure I.3. Newton thinking and NASA boffin computing.

> Furthermore, the simplicity of mathematics has long been held to be the key to applications in physics. Einstein is the most famous exponent of this belief. But even in mathematics itself the simplicity is remarkable, at least to me; the simplest algebraic equations, linear and quadratic, correspond to the simplest geometric entities, straight lines, circles, and conics. This makes analytic geometry possible in a practical way. How can it be that simple mathematics, being after all a product of the human mind, can be so remarkably useful in so many widely different situations? Hamming (1980)

We can visualise what Hamming is talking about in Figure I.3. There we have Newton using the ideas of geometry to elaborate his "Mathematical Principles of Natural Philosophy". These are pure thoughts, but, remarkably, the mathematics mirrors what is happening in the solar system. There seems an uncanny correspondence between some simple principles of geometry in one man's mind and the movements of planets millions of miles away.

But were they just thoughts?

Or were they always something material, produced by physical work using physical tools?

Recall how Blake depicted Newton in Figure as a draftsman. Of course there are all sorts of Masonic overtones to this, echoes of the Great Architect etc, but it grasps a reality.

To explain his ideas to others, Newton had to resort to pictures, diagrams and arguments on the printed pages of his great book. The book itself was material, and copies survive to this day. While the thinking which went into writing the book was fleeting and perished along with the man, the maths seems eternal, independent of the man and of his paper book. It is as able to predict heavenly motions now as it was in the 17th century.

Figure I.4. Blake's Newton.

But now look at bottom of Figure I.3. It depicts the early days of space exploration. A boffin sits at his old valve computer and gets it to work out the course that will take a probe to Mars.

At one level this shows the same process as the top illustration, but the very physicality of that big grey metal box in front of the boffin hints at something different. The similarity is that Newtons laws, and their associated mathematics are being used in each case. But the fact that the calculations are now taking place in a machine makes it harder to see the process as being one of a correspondence between mathematical thought and reality.

To do his calculations the NASA scientist would have had to have fed the computer with a whole lot of data obtained by astronomers. He will have had to develop programs to represent the dynamics in question. And he then set it off working. We say he will have had to develop programs, but, that is not strictly necessary. The computer in the picture is actually an analogue one, which was programmed by rewiring it using the patch panel behind the operator.

So the correspondence here is actually one between the dynamics one physical system – the analogue computer, and the dynamics of another – the rocket that the designer wishes to control. This idea that a mathematical model is actually a physical thing, seems very modern, but is in fact ancient. Before Newton came Kepler (Fig. I.5), before him Ptolemy. Before Ptolemy came Hipparchus and Apollonius (Fig. I.6). Ptolemy's epi-cycle model is well known, but it is equivalent to Apollonius's Cycle and Deferent Model (Fig. I.7).

a) Hipparchus's actual model? : In 1900 a group of sponge divers sheltering from a storm anchored off the island of Antikythera. Diving from there they spotted an ancient shipwreck with bronze and marble statuary visible. Further diving in 1902 revealed what appeared to be gearwheels embedded in rock. On recovery these were found to be parts of a complicated mechanism, initially assumed to be a clock. Starting in the 1950s and going on to the 1970s the work of Price established that it was not a clock but some form of calendrical computer. Using X-rays, modern reconstructions have been built showing that it physically implemented Apollonius's model of the lunar orbit. The original machine dates

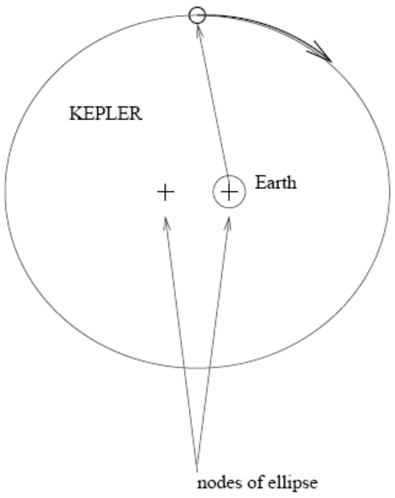

Figure I.5. Kepler's model of lunar motion

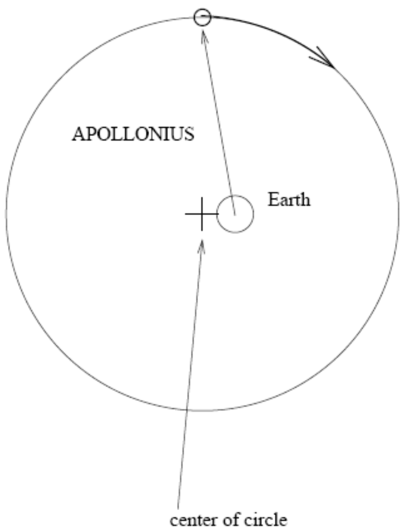

Figure I.6. Apollonius's model of lunar motion

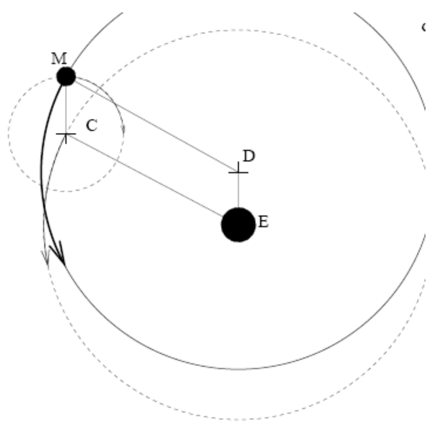

Figure I.7. Equivalence of Ptolemy and Apollonius models. Bold lines indicate eccentric model dashed lines the epi–cycle model.

Figure I.8. Tania van Vark's beautiful reconstruction of the Antikythera computer.

from the 2nd century BC but modern reconstructions have been built. I show a particularly beautiful one by Tania van Vark. You turn the handle and get predictions of the position of the sun and moon in the sky and the dates of eclipses. It emphasizes how a scientific model is a microcosm emulating a macrocosm.

Since the invention of the Universal Computer in the 1940s, it was no longer necessary to build special purpose mechanical models of physical system.

A universal computer is a physical device that can be configured to simulate any physical process. It is configured by the input of an appropriate mathematical function representing the model. Once that is done it becomes a physical device that models another physical process.

The Church-Turing-Deutsch (CTD) Principle, after the three people (Alonzo Church, Alan Turing, and David Deutsch) who contributed most to the formulation of the principle, states that

> *Every physical process can be simulated by a universal computing device.*

II. Modeling and Entropy

We can now summarize some key principles of modeling.

- A model is a physical subsystem, now often implemented on a universal computer.
- It is involved in the generation of testable predictions – A model which makes no testable predictions is useless
- We want it to display **elegance or simplicity** – Occam's Razor 'Entities should not be multiplied without cause'. This has long been an ideal for modelers and theorists, but the development of Chaitin Kolmogorov information

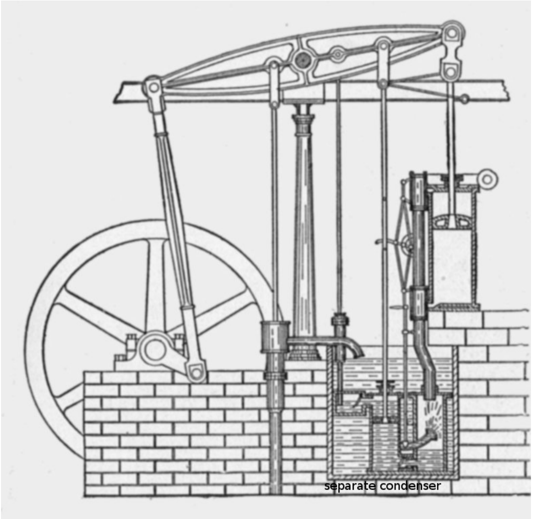

Figure II.1. James Watt worked on a model steam engine at Glasgow in 1765 from which his invention of the separate condenser came. This laid the foundations both for practical applications of steam power and also for the systematic study of the relation between heat and work from which the concept of entropy arose.

theoryChaitin (1999); Li and Vitanyi (1997) explains why it is so important.

Mathematically we can view any use of a model as $(p,d) \leftarrow M(d)$ Where p are the predictions, M is the function encoding the model, and d are the input data. After running the model we have both the predictions and the original data. For the model to be elegant we want to maximise its information yield $Y = I(p)/I(M)$

Where by $I(x)$ we mean the information content of x. For the model to be useful we want to maximise the mutual information $I(p;o)$ in the predictions p, and observed system o. That is to say we want to maximise the information that is common to the predictions of the model and the observations.

Max $I(p;o) = H(p)–H(p|o)$

Whilst

Min $I(M)$

minimising the information in the model $I(M)$. Where $H(p)$ is the uncertainty or entropy in p and $H(p|o)$ is the uncertainty in p given o That is to say we want to explain as much of the data as we can but we should avoid models that contain a lot of internal information – in the worst case such a model simply tabulates the observations and has no general predictive ability when fed with different data.

Why entropy?

In the formula to find mutual information we used the H function for entropy. Why?

Surely entropy has to do with thermodynamics which studies things like the efficiency steam engines?

Yes that is true. That is how thermodynamics originated, but a key discovery of the 20th century was how information and entropy are linked.

The Basic Problem of Information :

- What is information?
- How does it relate to entropy?

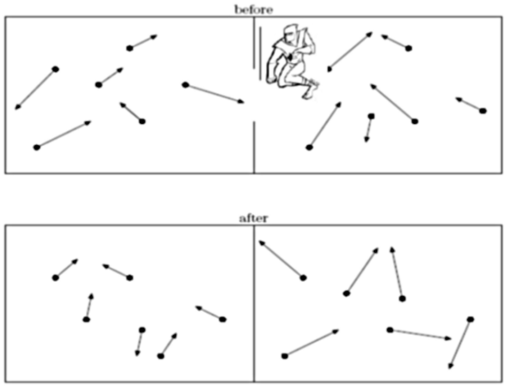

Figure II.2. Maxwell's Daemon Gas initially in equilibrium. Daemon opens door only for fast molecules to go from A to B, or slow ones from B to A. ! Slow molecules in A, fast in B. B hotter than A, and can be used for power. Information has produced power!

Clausius established that : no process possible that has the sole effect of transferring heat from a colder to a hotter body. This implied that, for instance, there was no chance of transferring the heat wasted in the condenser of a steam engine back to the boiler where it would boil more water. Thermodynamics ruled out perpetual motion machines. It was the first form in which the concept of thermodynamic irreversibility arose.

But James Clerk Maxwell, one of the early researchers in thermodynamics, came up with an interesting paradox.

> One of the best established facts of thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which temperature and pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. But if we can conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being would be able to do that which is presently impossible to us. For we have seen that the molecules in a vessel full of air at a uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without the expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics. (Maxwell (1875), pp. 328–329)

Boltzmann: Maxwell's proposed counter-example to the second law was explicitly based on atomism. With Boltzmann

(1995), entropy is placed on an explicitly atomistic foundation, in terms of a sum over molecular distributions in phase space.

$$S = -kN\sum p_i \log p_i$$

where i indexes over volume in six-dimensional phase space, p_i is the function that counts the probability of molecules being present in that volume, and k is Boltzmann's constant, and N the number of molecules. We will take a simplified version of Boltzmann's concept to explain the idea. Consider a gas in a cylinder. If the gas is evenly spread through the cylinder the entropy will be higher than if the gas is all concentrated at one side of the cylinder. Suppose we divide the length of a 10cm cylinder into 100 sub-lengths each of 1mm, and let us suppose that there are N molecules in the whole cylinder.

If the molecules are evenly divided along the length, each 1mm length the probability of a molecule being in any one section is 1/100, and the entropy will be proportional to $-K_{10}N \times 100/100 \times \log(1/100)$ which assuming we use $\log_{10}$instead of natural logs will give us a value of $2NK_{10}$ (K having been adjusted from Boltzmann's constant to allow for using base 10).

Now suppose that all of the gas molecules are concentrated, again evenly, in the first half of the cylinder where each 1mm length contains 20,000 molecules, with none in the rest. The entropy of the system is now proportional to $-K_{10}N \times 50/50 \times \log(1/50)$ or $1.69NK_{10}$ which is clearly less than the case were the gas was evenly spread. If we allow the concentrated gas to spread out, as happens in a pressure engine, then there is an increase in entropy.

Highly spread out distributions have a higher distribution and concentrated distributions have a low entropy.

Shannon : The communications engineer Shannon (1948) introduced the concept of entropy as being relevant to sending messages by teletype. The mean information content of an ensemble of messages is obtained by weighting the log of the probability of each message by the probability of that message.

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i$$

He showed that no encoding of messages in 1s and 0s could be shorter than this, this is essentially the same as Boltzmann's formula, expressed on a discrete basis.

Hence *information = entropy*.

Information measured in bits provides a common means of measuring both a model $M$and the predictions p

Suppose we have a vector of observations O which we have reason to believe are given to an accuracy of 3 digits. Then each observation contains $Log_2(1000)bits \approx$ 10bits

Suppose we have a prediction vector P which we assume is to the same accuracy. We can estimate $H(P|O)$ by histograming the distribution of the ratio P/O and then applying Shannon's entropy formula to the distribution. The information content of the model itself can also be estimated.

If we want to compare models, we can decompose each of the models into two parts

- A basic structure or formula
- A set of auxiliary parameters or constants that has to be provided

Each of these can be given an information measure. The formula is measurable in terms of the number of bits needed to write it down as a string of digital characters. The parameters are measurable in terms of the number of parameters and the accuracy in bits to which each has to be given.

A. Models and Laws

Sciences designate as laws those models that:

1) Have a simple, elegant formulation with few parameters
2) Make excellent predictions in an apparently unlimited number of cases

Applying this to economics one may ask how much of what is taught in undergraduate economics consists of

- Empirically testable and empirically tested propositions
- Formulae that are elegant and simple
- Simple formulae that are so universal and excellent in their predictive power as to deserve the name Laws.

There is obviously an immense wealth of empirical studies in the economics literature. But I am more concerned here with the basic theory that is taught to students starting economics. I am concerned with what these students would be learning when a physics student would be learning classical mechanics, or when an electrical engineer would be learning Maxwell's equations.

How well founded are the models that the economics students are taught?

What is the record of testability of neoclassical subjective value theory?

I am writing as an outsider. I only studied neo-classical economics to undergraduate level. Whereas in my Physics, Psychology or Biology courses we were given accounts of classic experiments that had verified for instance the invariability of the speed of light, the mechanism of reflex conditioning, or the germ theory of disease, our economics lecturers cited no such empirical studies when discussing basic theory.

When I was a student my economics professor told us that whilst the labour theory of value had been an important historical stage in the development of economics, it was now known to be fatally flawed. At that time I was told that the labour theory of value was now known to be inaccurate and superseded by the subjective utility theory. 20th century economists such as Sraffa and Samuelson had shown that it was unnecessary to accord labour any special place in our understanding of prices. Instead, the structure of prices could be perfectly well understood as the result of the monetary costs faced by firms and the behaviour of profit maximising entrepreneurs. If there was in reality no such thing as labour value, it followed that Marx's theory of exploitation was an invalid incursion of moral prejudices into the 'positive science' of economics.

The professor who taught us this, Ian Steedman, was actually quite left wing, an active member of the Communist Party. This is just an anecdote, but fact that even a prominent communist intellectual believed that the central component of Marx's theory was scientifically worthless is significant.

It shows how strong the intellectual dominance of orthodox economics had become.

But it is hard to see how the subjective theory is even testable. What quantitatively testable predictions about the price structure of the whole economy can it make?

Can one derive from it a predicted price vector to compare with the actual price vector?

If not, one has to put it in a basket labeled 'not even wrong'.

Testability – classical value theory : The classical theory of value on the other hand does make testable predictions. Once can make concrete predictions about market prices, using either Sraffa's formula

$$p = (1+r)Ap + w$$

Or the formula from Marx's Capital

$$v = Av + \lambda$$

where A is the technology matrix, p the price vector, v the value vector and λ the labour input vector, w the wage vector.

And you can then see how good the predictions each one makes are.

Until the 1980s it has to be said that the proponents of classical theories were as axiomatic in their approach as the proponents of neo-classical theories. Since then however there has been a growing realisation that not only were classical theories testable, but that they should be rigorously tested if any progress was to be made in resolving theoretical disputes. It was found, somewhat to the surprise of everyone, that the two leading classical theories gave almost the same accuracy in their predictions (Shaikh (1998); Cockshott and Cottrell (1997); Ochoa (1989); Tsoulfidis and Maniatis (2002)), in which case parsimony may favour the simpler model. I hope that later speakers will touch on this showing in practice how to the labour theory of value can be tested using an information theoretic measure.

Statistical mechanics and value.: 27 years ago two mathematicians Moshe Machover and Emanuel Farjoun, wrote a book called the Laws of Chaos(Farjoun and Machover, 1983). Their book gave a radically new way of looking at how capitalism worked as a chaotic and disorganised system. Farjoun and Machover had the the insight to see that physics had already developed theories to describe similar disorganised and chaotic systems.

In a market economy, hundreds of thousands of firms and individuals interact, buying and selling goods and services. This is similar to a gas in which very large numbers of molecules interact, bouncing off one another. Physics speaks of such systems as having a 'high degree of freedom', by which it means that the movements of all individual molecules are 'free' or random. But despite the individual molecules being free to move, we can still say things about them in the aggregate. We can say what their average speed will be (their temperature) and what their likely distributions in space will be.

The branch of physics which studies this is statistical mechanics or thermodynamics. Instead of making deterministic statements, it deals with probabilities and averages, but it still comes up with fundamental laws, the laws of thermodynamics, which have been found to govern the behaviour of our universe.

Now here is the surprise! When they applied the method of statistical mechanics to the capitalist economy, they found that the predictions it made coincided almost exactly with the labour theory of value as set out in volume 1 of Marx's Capital. Statistical mechanics predicted that the selling prices of goods would vary in proportion to their labour content just as Marx had assumed. Because the market is chaotic, individual prices would not be exactly equal to labour values, but they would cluster very closely around labour values. Whilst in Capital I the labour theory of value is just taken as an empirically valid rule of thumb. Marx knew it was right, but did not say why. Here at last was a sound physical theory explaining it.

It is the job of science to uncover causal mechanisms. Once it has done this it can make predictions which can be tested. If two competing theories make different predictions about reality, we can by observation determine which theory is right. This is the normal scientific method.

Farjoun and Machover's theory made certain predictions which went directly against the predictions made by critics of Marx such as Samuelson. In particular their theory predicts that industries with a high labour to capital ratio will be more profitable. Conventional economics predicts that there will be no such systematic difference between the profit rates in different industries. When put to the test it turned out that Farjoun and Machover were right. Industries with a high labour to capital ratio are more profitable(Zachariah (2006); Cockshott and Cottrell (1998)). But this is exactly what we should expect if the source of profit was the exploitation of labour rather than capital. Their theory made predictions which not only turned out to be empirically spot on, but at the same time verified Marx's theory of the exploitation of the worker.

III. Idea of a conservation law.

I said that the most powerful models for the predicting of reality are called laws. Among the most paradigmatic here are conservation laws. In physics conservation of energy, charge, probability.

Conservation laws reveal hidden symmetries in the structure of reality. This was discovered by Emily Noether in the flurry of work that followed the publication by Einstein of relativity theory.

Noether's theorem states that *any symmetry of the action of a system has a corresponding conservation law*.

The symmetry and the conservation are not necessarily immediately evident as the following examples will show. Put very loosely, Emmy Noether's principle says that in a physical system, a conserved quantity at one level of abstraction corresponds to a symmetry property of the system at another level of abstraction.

– Translational symmetry implies conservation of momentum. For a simple proof of this see the web page Noethers theorem in a Nutshell..

– Temporal symmetry implies conservation of energy

In order to explain these ideas I will take a couple of examples from physics before moving to show how the ideas

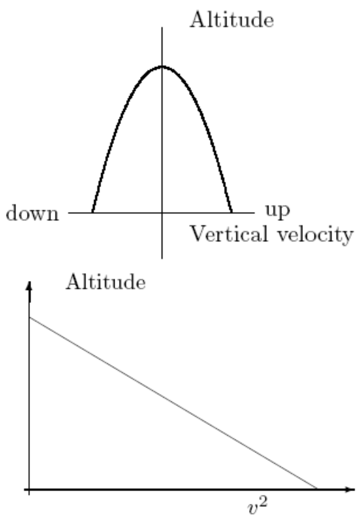

Figure III.1. Plots for a ball thrown up. Top shows points in phase space traversed by a projectile thrown upward in a gravitational field. Bottom, points in the space of (altitude, velocity squared) traversed by the particle.

of symmetry and conservation can be applied in the study of economic phenomena. Let us look first at the conservation of energy.

Figure III.1 shows two different ways of plotting the motion of a ball thrown upwards. The top diagram shows the vertical velocity plotted against height. The plot is parabolic and looks rather uninformative until we transform it to the bottom diagram which shows conservation of energy. Here we have height as one axis and the square of vertical velocity as the other. The footprint of a conservation law is shown the the straight line relationship between height (potential energy) and velocity squared (kinetic energy). A straight line corresponds to the equation $C = Ax + By$ with C the conserved quantity. Note that although we take energy for granted as a 'thing', it is in fact only revealed to us as something existing by these conservation laws.

We have the conservation law, now find the symmetry.

If we transform to yet another representation, in this case by taking the square root of the axes[1] we find the path is a circle (Figure III.2). In this representation the plot shows rotational symmetry. A projectile's movement in this representation constitutes a rotation that preserves distance from the origin.

Here we can see the symmetry associated with the conservation of energy. The new representation means that we can treat v the path as the result of rotational symmetry in a vector space.

Another example is quantum mechanics. Consider a simple example like photon polarisation. Two polarisers at right angles completely block the passage of photons. But if we pass light through a horizontal polariser and then measure the polarisation of the photons using a polariser at angle θ we find

[1]One first has to scale the 2nd graph in Figure III.1 in uniform units of energy to give a 45°slope.

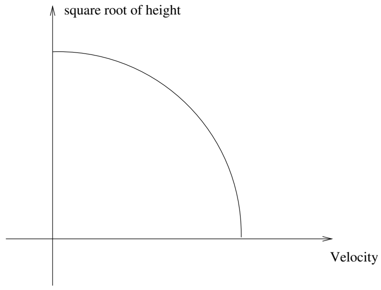

Figure III.2. If we plot v against $\sqrt[2]{h}$ we get a circle. Here the conserved quantity is distance from origin, and the system displays rotational symmetry in this coordinate system.

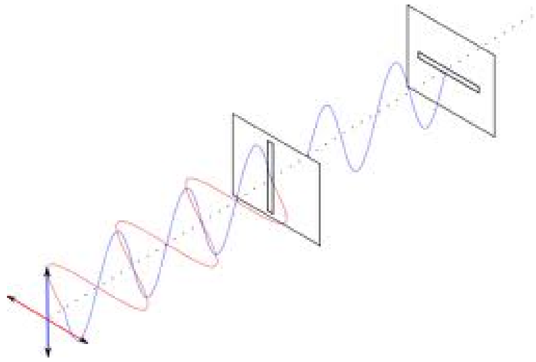

Figure III.3. Polarisation, orthogonal polarisers block transmission..

that the number of photons going through the second polariser varies as we change θ. We explain this by saying that in the frame of reference of a polariser we can treat the photons as being superpositions of two basis states $\alpha \mid horiz\rangle + \beta \mid vert\rangle$ with the variables $\alpha\beta$ being called amplitudes. The two basis states are labeled here with the orthogonal polarisation directions. We find that the probability of observing a photon in a given polarisation direction is the square of its amplitude. The amplitudes follow the relation $\alpha = \sin\theta, \beta = \cos\theta$. In this case of course the rotational symmetry of the maths is clear and can be intuitively understood in terms of the rotation of the polarisers. But the basic maths of amplitudes is much more general than this particular example and is used throughout quantum mechanics. Why do quantum physicists use amplitudes whose squares gives us probabilities?

It is surely because this vector space, which allows the application of unitary rotation operators, projects onto the space of probabilities which are a scalar conserved property: probabilities always sum to 1. This is the same as the change of representation we had to illustrate conservation of energy.

Similarity between value and energy metrics : Commodity money space is not a vector space, since the metric it follows is $d = |\alpha\Delta_x + \beta\Delta_y|$, these are captured by the isovals or budget lines shown in Figure III.4. We can draw such a diagram for any two commodities that are being traded.

This makes it analogous with a system with a hidden conserved quantity like energy. In this case the conserved quantity is what we call value. One can develop(Cockshott, 2005, 2009) the concept of an underlying space, commodity amplitude space, which can model commodity exchanges

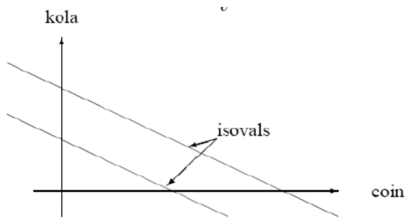

Figure III.4. Isovals in commodity money space. The isovals show the combinations of a commodity (Kola in this case) and coins that can be reached starting off from a particular number of coin.

and the formation of debt. Unlike commodity space itself, this space, is a true vector space whose evolution can be modeled by the application of unitary rotation operators. The relationship between commodity amplitude space and observed holdings of commodities by agents is analogous to that between amplitudes and observables in quantum theory.

I would argue that the application of Noether's theorem gives us a more rigorous way of formulating what Marx was arguing in the chapter on the forms of value in volume I of capital. These should be seen as an attempt to define commodity exchange and exchange value as a conservative system.

IV. Applying conservation laws and entropy to wealth distribution

Thermodynamics predicts that systems tend to settle into a state of maximum entropy. The conservation laws specify that whilst this randomization occurs energy must be conserved. Boltzmann and Gibbs showed that this implies that the probability distribution of energies E_i that meets these two criteria is

$$p_{E_i} = e^{\frac{E_i}{kT}}$$

YakovenkoDragulescu and Yakovenko (2002) has argued that since money is conserved in the buying and selling of commodities it is analogous to energy. If the system settles into a maximum entropy state then monetary wealth will come to follow a Gibbs Boltzmann distribution. He is able to show that the observed income distribution for 96% of the US population is well explained by a negative exponential distribution of the Gibbs form: Figure IV.2.

There remains a super-thermal tail of income (the top 4%) whose income is not conformant with maximal entropy but follows a power law distribution.

> The straight line on the log-linear scale in the inset of Fig. IV.2 demonstrates the exponential Boltzmann-Gibbs law, and the straight line on the log-log scale in the main panel illustrates the Pareto power law. The fact that income distribution consists of two distinct parts reveals the two-class structure of the American society . Coexistence of the exponential and power-law distributions is also known in plasma physics and astrophysics, where they are called the "thermal" and "super-thermal" parts . The boundary between the lower and upper classes can be defined as the intersection point of

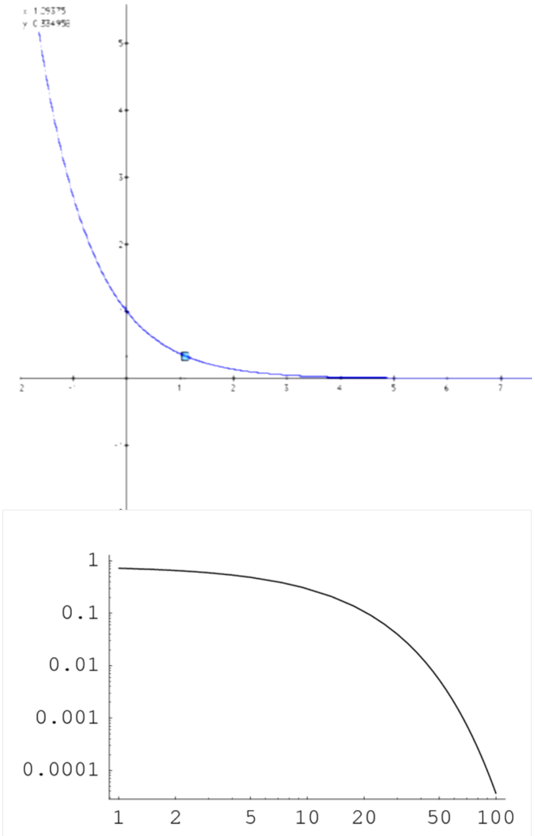

Figure IV.1. The Gibbs distribution, top plot on a linear scale, bottom on a log log scale.

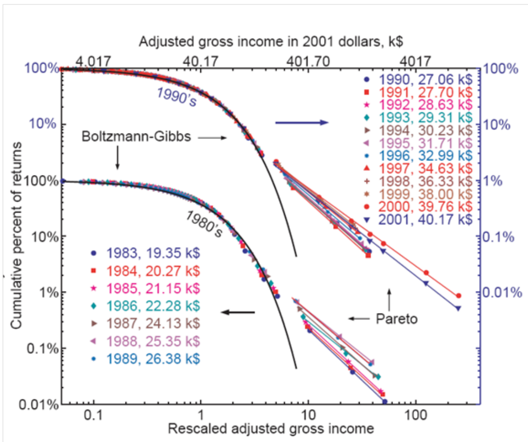

Figure IV.2. Yakovenko shows that the distribution of wealth follows basically a Gibbs distribution with a super thermal tail. Reproduced from Yakovenko and Rosser Jr (2009).

> the exponential and power-law fits in Fig. IV.2. For 1997, the annual income separating the two classes was about 120 k$. About 3% of the population belonged to the upper class, and 97% belonged to the lower class.(Yakovenko and Rosser Jr, 2009)

The thermal distribution arises from the application of the conservation law plus randomness. The non thermal distribution from the violation of conservation. Tied to income from capital and the stock market. This is consistent with Marx's analysis that profit in general can not arise within a conservative system, but from something outside of the conservative system – production of surplus value.

> The subject of income and wealth distributions and social inequality was very popular at the turn of another century and is associated with the names of Pareto, Lorenz, Gini, Gibrat, and Champernowne, among others. Following the work by Pareto, attention of researchers was primarily focused on the power laws. However, when physicists took a fresh look at the empirical data, they found a different, exponential law for the lower part of the distribution. Demonstration of the ubiquitous nature of the exponential distribution for money, wealth, and income is one of the new contributions produced by econophysics. The motivation, of course, came from the Boltzmann-Gibbs distribution in physics. Further studies revealed a more detailed picture of the two-class distribution in a society. Although social classes have been known in political economy since Karl Marx, realization that they are described by simple mathematical distributions is quite new. Very interesting work was done by the computer scientist Ian Wright (Wright, 2005, 2008), who demonstrated emergence of two classes in an agent-based simulation of initially equal agents.(Yakovenko and Rosser Jr, 2009)

Wright has shown that random exchange models generate combined Gibbs + power law distributions as soon as you allow the hiring of labour. This is again consistent with Marx's old analysis.

Summary

- A model must make testable predictions to be scientifically meaningful.
- Information theory gives us a uniform means to measure both models and the predictions of models in order to evaluate their adequacy.
- One should not be afraid to make use of it.
- The results of tests using modern methods are consistent with broadly Marxian models.

References

Boltzmann, L.: 1995, *Lectures on gas theory*. Dover Pubns.

Chaitin, G.: 1999, 'Information and Randomness: A survey of Algorithmic Information Theory'. In: *The Unknowable*. Signapore: Springer.

Cockshott, P.: 2005, 'Symmetries and exchange relations'. *Politica y Cultura* **23**, 3–32.

Cockshott, P.: 2009, 'Hilbert Space Models Commodity Exchanges'. *Quantum Interaction* pp. 299–307.

Cockshott, P. and A. Cottrell: 1997, 'Labour time versus alternative value bases: a research note'. *Cambridge Journal of Economics* **21**, 545–549.

Cockshott, P. and A. Cottrell: 1998, 'Does Marx need to transform?'. In: R. Bellofiore (ed.): *Marxian Economics: A Reappraisal*, Vol. 2. Basingstoke: pp. 70–85.

Dragulescu, A. and V. M. Yakovenko: 2002, 'Statistical mechanics of money, income and wealth: a short survey'. http://arXiv.org/abs/cond-mat/0211175.

Farjoun, E. and M. Machover: 1983, *Laws of Chaos, a Probabilistic Approach to Political Economy*. London: Verso.

Hamming, R.: 1980, 'The unreasonable effectiveness of mathematics'. *American Mathematical Monthly* pp. 81–90.

Li, M. and P. Vitanyi: 1997, *An Introduction to Kolmogorov Complexity and Its Applications*. Springer.

Maxwell, J. C.: 1875, *The Theory of Heat*. London: .

Ochoa, E. M.: 1989, 'Values, prices, and wage–profit curves in the US economy'. *Cambridge Journal of Economics* **13**, 413–29.

Shaikh, A. M.: 1998, 'The empirical strength of the labour theory of value'. In: R. Bellofiore (ed.): *Marxian Economics: A Reappraisal*, Vol. 2. Macmillan, pp. 225–251.

Shannon, C.: 1948, 'A Mathematical Theory of Communication'. *The Bell System Technical Journal* **27**, 379–423 and 623–56.

Tsoulfidis, L. and T. Maniatis: 2002, 'Values, prices of production and market prices: some more evidence fro the Greek economy'. *Cambridge Journal of Economics* **26**, 359–369.

Wigner, E.: 1960, 'The unreasonable effectiveness of mathematics in the natural sciences'. *Communications in Pure and Applied Mathematics* **13**(1), 1–14.

Wright, I.: 2005, 'The social architecture of capitalism'. *Physica A: Statistical Mechanics and its Applications* **346**(3-4), 589–620.

Wright, I.: 2008, 'Implicit microfoundations for macroeconomics'. *Economics Discussion Papers*.

Yakovenko, V. and J. Rosser Jr: 2009, 'Colloquium: Statistical mechanics of money, wealth, and income'. *Reviews of Modern Physics* **81**(4), 1703–1725.

Zachariah, D.: 2006, 'Labour Value and Equalisation of Profit Rates'. *Indian Development Review* **4**(1), 1–21.