Can Computers Do the Job of Nobelist Physicists? Planck Formula Revisited

Richard Aló[1], Raúl Trejo[2], and Vladik Kreinovich[3]

[1]Center for Computational Sciences and
Advanced Distributed Simulation
University of Houston–Downtown
Houston, TX 77002, USA, RAlo@uh.edu

[2]Sistemas de Informacion
ITESM Campus Estado de México
Apdo. Postal 2, Modulo de Servicio Postal
Atizapan, México 52926
rtrejo@campus.cem.itesm.mx

[3]Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA, vladik@cs.utep.edu

Abstract

There exist several computer programs which successfully model the discovery process in science. There are successful expert systems in medicine and other areas. But one area is a real challenge for such systems: theoretical physics. The most advanced knowledge discovery programs (like BACON written under the supervision of the Nobelist Herbert A. Simon) successfully reproduce only 17, 18, and 19 century physics, but stop short of explaining the very first formula of the 20 century: Planck's law of black body radiation. This law, discovered by an insight, led to the modern Quantum Physics. The programs stop short not because the computers are not fast enough: as Simon emphasized, we need new ideas – not only new computers.

In the present paper, we present the natural symmetry ideas which lead directly to Planck's formula. Possible other applications of these ideas are discussed.

1 A Challenge of Physics

Expert systems are efficiently used by medical doctors, by engineers, in many other spheres, but (alas!) theoretical physics is still out of their scope. We have programs that come to the same (or even better) conclusions about diagnoses and possible treatment as doctors but we have no programs that can come to the same conclusions as physicists.

One can hide behind the results like that of [2, 3] that physical problems are NP-hard or in some other sense undecidable, but in many other cases there are programs that efficiently solve some instances of the problems that are in general undecidable, so what's wrong here?

OK, maybe modern physics is very complicated but the problem is that we cannot even get the formulas that the physicists have discovered long ago. The most advanced system especially aimed at computer modeling of scientific discovery, the BACON system [5] managed to deduce from data the main laws of 17, 18 and even 19 century physics, but failed to reconstruct historically the first law of 20 century physics: Planck's 1900 law of black-body radiation. As Simon stresses in his paper, this is not because we ran out of time or space: but because the system was badly in need of a new idea. The ideas of monotonicity and conservation laws that were ingeniously incorporated into the BACON system did not work in this case. This means that we are missing an idea, and so maybe adding this idea will help? We present a new idea, that enables to deduce Planck's formula. Crudely speaking, this idea combines symmetry groups theory (originally a typical physical idea, but with lots of nontrivial applications to computer science [4]) with algebraic complexity (from the computer science side).

2 Why Was Planck's Formula So Important?

The 19 century physics failed to explain the radiation of an absolutely black body: it predicted that the energy density $E(\nu)$ of radiation must increase as a square of frequency $C \cdot \nu^2$. This formula was good for small ν, but failed for large ones. For large ν, an empirical formula $\nu^3 \cdot \exp(-\nu/kT)$ was discovered, but this formula, in its turn, did not work for small ν. So we need to find an expression that would work for all ν and have given asymptotics for ν tending to 0 and to ∞. In principle, there exist lots of such functions, so the problem of choice at first glance must be resolved experimentally. But Planck managed to find the expression, namely, $\nu^3/(\exp(\nu/kT) - 1)$, which is not only experimentally confirmed by all the experimental data of the last 100 years, but actually formed a foundation of new physics: quantum mechanics.

It is not the problem that the physicists know how this formula was discovered, and we cannot reproduce that sequence of steps in a computer. The problem is worse: the physicists cannot explain that discovery in understandable terms.

3 Group Symmetry Idea

In order to understand Planck's formula let's analyze the simpler cases, when the path to discovery is known (see, e.g., [1]).

For example, for many astrophysical sources the dependence of energy density on wavelength is described by a simpler formula $E(\nu) = A \cdot \nu^{-\alpha}$, where A and α are constants. Feynman explains it in such a way: of course, the real processes are not scale invariant in the sense that some processes have fix spatial sizes, so when we describe them in cm or in meters, we get completely different formulas. However, when the frequency is sufficiently big, we can consider these fixed sizes to be infinitely big and thus our expressions to be invariant with respect to the change of length unit ($\nu \to \lambda \cdot \nu$, where λ is a ratio of those units). This invariance demand in mathematical terms is $E(\lambda \cdot \nu) = C(\lambda) \cdot E(\nu)$, where C is a constant; one can mathematically prove that the only solutions of this equation are $A \cdot \nu^{-\alpha}$ (see, e.g., [4]).

Other simple formulas can be obtained from different symmetry considerations. Our example is energy distribution in astrophysical sources, i.e., asymptotic formulas describing the dependence of $N(E)$ on E, where $N(E) \cdot dE$ is the number of particles per unit volume with energies from E to $E + dE$.

Strictly speaking, in general case, there are no symmetries between different values of energy, because there is always a fixed value of energy $E_0 = m_0 \cdot c^2$, corresponding to the rest mass m_0 of the particles which form this astrophysical object. However, when $E \gg E_0$, we can ignore the rest mass in comparison with the total particle energy, and with great accuracy, we can view these particles as massless. For massless particles, there is no fixed energy, therefore equations, describing their distribution, must be invariant with respect to the change of the unit in which energy is measured. If we change energy units to λ times smaller ones then numerical values (measured w.r. to those units) increase λ times: $E \to \lambda \cdot E$. Of course, if we change the unit of energy, then, for the equations to remain valid, we must change the other units accordingly.

In particular, a unit of volume can change. Therefore, in the new units, the distribution of particles per volume and per energy is described by the formula $C \cdot N(\lambda \cdot E)$, where C is a constant. Since we have the exact same distribution but expressed in two different units, we must therefore have $C \cdot N(\lambda \cdot E) = N(E)$, hence so $N(\lambda \cdot E) = c \cdot N(E)$ for some constant c.

One can easily prove that the only continuous (or even measurable) functions satisfying this property have the form $N(E) = A \cdot E^a$: indeed, from $N(\lambda \cdot E) = c(\lambda) \cdot N(E) = c(E) \cdot N(\lambda)$, we conclude that $N(E)/c(E)$ is a constant. Dividing both sides of the original equation by this constant, we conclude that $c(\lambda \cdot E) = c(\lambda) \cdot c(E)$ hence $c(E) = E^a$. Therefore, $N(E) = A \cdot E^a$ for some constant A.

Indeed, asymptotics of the type $N(E) \sim A \cdot E^a$ is typical for hot matter energy distributions (i.e. for large E).

In the opposite case of small $E (E \to 0)$, the situation is quite different: $E << m_0 \cdot c^2$, therefore we can assume that $m_o \cdot c^2 \approx \infty$. Here there is no absolute initial point for measuring energy, therefore the energy distribution must be invariant with respect to energy shift, i.e. $N(E + E_0) = c(E_0) \cdot N(E)$. From this equation one can conclude that $c(E_0 + E_1) = c(E_0) \cdot c(E_1)$ hence $\ln(c(E))$ is an additive function, so $\ln(c(E)) = k \cdot E$ and $c(E) = \exp(-k \cdot E)$, $N(E) = A \cdot \exp(-k \cdot E)$. This is a typical energy distribution for nonrelativistic case (small E) – the well-known Gibbs distribution.

These examples motivate the following approach.

4 Main Idea

Although in reality processes are not precisely symmetric, but there are often approximate symmetries that are the more precise the closer we come to some critical value of some parameter x. In case we neglect symmetry violations we obtain symmetric expressions which can be called basic asymptotics.

Real-life processes can be more complicated because each process is influenced by several factors. Even when each factor is described by a single asymptotic, their combination makes the result more complicated. This joint effect can be described as a combination of basic asymptotics.

Which combinations shall we use? The combination is an operation transforming functions into functions $f(x),\ g(x) \to h(x)$. First of all it seems natural to demand that this combination rule be *local*, i.e., the value of $h(x)$ should depend only on the values of $f(x)$ and $g(x)$ only in the same point x, i.e. $h(x) = f(x) * g(x)$, where $*$ is some real function of 2 variables. (Simplest case – linear superposition: the field of two charged particles is the sum of the fields, induced by each of them, here $* = +$.) Due to this "composition" idea this operation must satisfy the natural demands like $a * b = b * a$ and $(a * b) * c = a * (b * c)$ (it should not matter in what order we combine these solutions); so, it must be a commutative semigroup. Moreover, it is natural to assume that the combination result really depends on both inputs, i.e., that if $b \neq c$ then $a * b \neq a * c$. As a result, we conclude that $a * b = a * c$ implies $b = c$. Therefore, $*$ can be extended to a commutative group operation.

The main reason why the combination is necessary is that different combined processes have different symmetries. Therefore it is natural to require that when the combined processes have the same symmetry then the combination $a * b$ should be of the same symmetry itself.

We show that these demands lead to $* = +$ (or $a * b = (a^p + b^p)^{1/p}$) or $* =$ product.

Then the natural partial ordering meaning "A is less complicated than B" is introduced (for example, basic functions are less complicated than their combinations; combination using three operations $*_1, *_2, *_3$ is more complicated than the one using its subset $\{*_1, *_2\}$ etc.), and we prove that Planck formula is indeed simplest in this sense.

Let's now turn to formal definitions.

5 Main Definitions and Results

In the following text, by a *function* we understand a continuous function $f : R \to R$ from real numbers to real numbers.

Definition 1.

- *We say that a function f is shift-invariant if for every x_0 there exists a constant $c(x_0)$ such that $f(x + x_0) = c(x_0) \cdot f(x)$.*

- *We say that a function f is scale-invariant if if for every $\lambda > O$ there exist a constant $c(\lambda)$ such that $f(\lambda \cdot x) = c(\lambda) \cdot f(x)$.*

We'll say that a function is *basic* if it is either shift-invariant or scale-invariant.

Proposition (see, e.g., [4]) **.** *A function f is shift-invariant if and only if it has the form $f(x) = A \cdot \exp(-k \cdot x)$; a function is scale-invariant iff it has the form $f(x) = A \cdot x^a$.*

Corollary. *The only functions that are both shift- and scale-invariant are constants.*

Comment. We can express this result by saying that constants are maximally symmetric hence they are the simplest basic functions. So we arrive at the following definition.

Definition 2. *We say the $f = const$ is simpler than any non-constant basic function f', and denote it by $f < f'$.*

Comment. Assume $*$ is a commutative group operation on R. It is well known that all such operations are of the type $x * y = \psi^{-1}(\psi(x) + \psi(y))$ for some function ψ, so we assume that $x * y = \psi^{-1}(\psi(x) + \psi(y))$.

Definition 3. *We say that two basic functions f, g have the same symmetry if either they are both shift-invariant with one and the same function $c(x_0)$, or they are both scale-invariant with one and the same function $c(\lambda)$.*

Proposition 1. *If $*$ is such that whenever $f(x)$ and $g(x)$ have the same symmetry then $h(x) \stackrel{\text{def}}{=} f(x) * g(x)$ has the same symmetry then $* = +$ or $a * b = (a^p + b^p)^{1/p}$*

Comment. This is the most symmetry preserving combination rule (therefore the simplest one – because the simplest case is also the most symmetric case). We can also consider less symmetric combination rules.

Proposition 2. *If $*$ is such that whenever $f_1(x), \ldots, f_n(x)$ are shift-invariant and have the same symmetry then $f_1 * \ldots * f_n$ is also shift-invariant (but not always have the same symmetry as each of f_i) then $*$ =product.*

(The same is true for scale-invariance.)

Comments. So all operations are either $(a^p + b^p)^{1/p}$, or $a \cdot b$. Turning if necessary to new units $a \to a^p$ (e.g. from length to area, etc.) we can now conclude that possible combining operations are $+$ and $\cdot$ (correspondingly, $-$ and :).

So, an arbitrary asymptotic expression can be obtained from basic ones by addition and multiplication, i.e., in mathematical terms, it is a *rational expression* in basic functions.

Remark 1. Another way to restrict ourselves to $p = 1$ is to assume that the operation $*$ is an analytical function at 0 (from physical viewpoint, it is a very natural demand).

Remark 2. One can also consider *unary* operations, i.e. mappings $\phi : R \to R$. Here likewise propositions are true.

Proposition 3. *If ϕ is such that for every basic $f(x)$ the function $\phi(f(x))$ has the same symmetry then $\phi(x) = A \cdot x$.*

Proposition 4. *If ϕ is such that for every basic f the function $\phi(f(x))$ is also basic, then $\phi(x) = A \cdot x^p$ for some A, p.*

6 Complexity Considerations

We have already mentioned that most symmetric is simplest, all complexity is due to symmetry violations. Therefore constants are simpler than other basic functions, $+(-)$ is simpler (because it preserves more symmetry than $\cdot(:)$). So, we arrive at the following definitions:

Definition 4. *By an asymptotic expression we mean arbitrary formula P obtained from basic functions by using $+(-)$ and $\cdot(:)$.*

Denotations. In this paper, we will denote asymptotic expressions by $P, Q, \ldots$ By $\mathrm{add}(P)$, we mean the total number of additions and subtractions in P, by $\mathrm{mult}(P)$ – the total number of multiplications in P, by $\mathrm{const}(P)$ – the total number of constants among basic functions used in P.

Definition 5. *We say that P and Q are of the same structure if $P(x) = R(e_1(x), \ldots, e_n(x))$ and $Q(x) = R(f_1(x), \ldots, f_n(x))$, where R is one and the same rational expression and e_i, f_i – basic functions. For each i we say that e_i and f_i are correspondent basic functions.*

Definition 6. *We say that a constant f is simpler that any other basic function g (and denote it by $f < g$), and that if f, g are both constant, or both non-constant exponents, or both non-constant degrees, then they are of same complexity ($f \sim g$).*

Definition 7. *We say that P is simpler or of same complexity as Q (and denote it $P \leq Q$) if add $P \leq$ add Q, mult $P \leq$ mult Q and one of the following 3 properties are true: either add $P <$ add Q, or mult $P <$ mult Q, or P and Q are of the same structure and corresponding basic functions satisfy $e_i \lesssim f_i$ for all i.*

Definition 8. *We say that an expression P is simplest in the class $\mathcal{P}$ of expressions if whatever $Q \in \mathcal{P}$ we take either $P < Q$, or $P \sim Q$.*

Comment. Our definition of $P \lesssim Q$ seems rather restrictive. For example, it can seem reasonable to say that $P < Q$ if add $P =$ add Q, mult $P = \mathrm{mult}\, Q$ and const $P > \mathrm{const}\, Q$. But remember that our main objective is to prove that the Planck formula is the simplest possible one. The weaker definition of "simple" we take, the stronger the theorem. So we prove the theorem for the weakest notion of complexity – namely, for the the above one – bearing in mind that as soon as the theorem is proved, we can strengthen this definition as we wish – and the theorem will still remain true.

Let's now turn to the precise formulation of main theorem (its proof is in the appendix).

Theorem. *Among all asymptotic expressions which are asymptotically equivalent to $\nu^3 e^{-\nu/kT}$ for $\nu \to \infty$ and to $C\nu^2$ for $\nu \to 0$ there is a unique simplest expression, and it is the Planck formula.*

Comments.

- Of course, Planck formula is very simple, and one can always choose weights of algebraic operations so that from the viewpoint of the correspondingly weighted algebraic complexity Planck formula is the simplest. What we prove is not that trivial result. We prove that, crudely speaking, whatever weights we take, Planck formula is always the simplest – so it is really simplest ("most symmetric").

- We spent some time to prove that the operations are $+$ and $\cdot$ are the simplest. What if we supposed that from the very beginning? These are the natural operations, normally used in algebraic complexity considerations. Of course we could do so; but then the result that Planck formula is the simplest would be quite natural from a computer scientist viewpoint, but completely mystical from the physicists' one: indeed, why should real world (and what we wish to describe in physics is namely real world) be constructed in such a way that it simplifies the calculations connected with it? It's hard to think of examples in favor of this strange hypothesis; moreover, the extremely complicated formulas of modern physics seem to disprove this hypothesis. On the other hand, if we demand that nature be symmetric or close to symmetric—this demand is (from physicists viewpoint) quite natural, because group symmetry considerations form one of the most productive ideologies of modern] physics. So our justifications of why $+$ and $\cdot$ are the simplest were not in vain.

- In many cases the so-called 2-point Padè approximation is useful, when one uses known asymptotics for $x \to 0$ and $x \to \infty$ in order to reconstruct the function for all x as a rational function of minimal possible degree. However this approach is applicable only when known asymptotics are of the type $A \cdot x^k$ with integer k. So our approach can be considered as a generalization of Padè ideology to arbitrary functions. For example, often $f(x) \sim A \cdot x^\alpha$ for $x \to 0$ and $\sim B \cdot x^\beta$ for $x \to \infty$, but α, β are not integers. Here if $\alpha < \beta$ then (as one can easily see) the simplest possible expression with these asymptotics is $Ax^\alpha + Bx^\beta$ (one cannot do without addition at all). If $\alpha > \beta-$ addition is not sufficient, and one can prove that the simplest is

$$\frac{1}{A^{-1}x^{-\alpha} + B^{-1}x^{-\beta}}$$

(1 addition, 1 multiplication).

7 Open Problems

- We defined only a partial ordering on the set of all possible asymptotic expressions. How to extend it to a total ordering? For example, if in one expression 3 multiplications and 2 additions are used, and in another 2 multiplications, and 5 additions – which expression is simpler?

  Is it possible to define this total complexity ordering relation in such a way that for arbitrary given asymptotics at finite many points (finite or infinite) there is a unique simplest expression with these very asymptotics?

- In Planck formula example we had no problem with finding at least one expression with given asymptotics – it is already known.

  But in general case – is there an algorithm allowing to give such as expression? Or, in view of the first open problem – an efficient algorithm (and not just enumerating all possible expressions) giving the simplest expression? Of course if we aim at AI we need algorithms not theorems.

- We considered only the case when basic functions are x^α and $\exp x$, because this is the case sufficient for Planck formula. However, in other situations there can be other basic asymptotics, e.g. sin or cos (corresponding to complex symmetries), log (inverse function to exp) etc. Our analysis of possible operations should be extended to these cases also.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209 and grant NCC 2-1232, by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-00-1-0365, by Grant No. W-00016 from the U.S.-Czech Science and Technology Joint Fund, and by Grant NSF 9710940 Mexico/Conacyt.

References

[1] R. Feynman, *Statistical mechanics*, 1973.

[2] V. Kreinovich, "Spacetime isomorphism problem is intractable (NP-hard)", *Int. J. Theor. Phys.*, 1991, Vol. 30, No. 9, pp. 1249–1257.

[3] V. Kreinovich, A. Vazquez, and O. Kosheleva, "Prediction problem in quantum mechanics is intractable (NP-hard)", *Int. J. Theor. Phys.*, 1991, Vol. 30, No. 2, pp. 113–122.

[4] H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.

[5] H. A. Simon, "Computer modeling of scientific and mathematical discovery processes", *Bull. Amer. Math. Soc.*, 1984, Vol. 11, No. 2, pp. 247–262.

Appendix: Proofs

Proof of Proposition 1

It's easy to see that shift-invariant functions have the same symmetry iff they are $A\cdot\exp(-k\cdot x)$ and $B\cdot\exp(-k\cdot x)$ for the same k. So the demand on $*$ is that for some C:

$$A\exp(-k\cdot x) * B\exp(-kx) = C\exp(-kx)$$

for all x. For $x=0$ we have $A*B=C$, therefore

$$A\exp(-kx) * B\exp(-kx) = (A*B)\exp(-kx).$$

But $t = \exp(-kx)$ is an arbitrary positive real, so $tA * tB = t(A * B)$ for all t, A and B. If we substitute $A * B = \psi^{-1}(\psi(A) + \psi(B))$ we obtain that if $\psi(A) + \psi(B) = \psi(C)$ then $\psi(tA) + \psi(tB) = \psi(tC)$ Likewise if

$$\underbrace{\psi(A) + \ldots + \psi(A)}_{n \text{ times}} = \underbrace{\psi(B) + \ldots + \psi(B)}_{m \text{ times}}$$

i.e. $n\psi(A) = m\psi(B)$ then $n\psi(tA) = \psi(tB)$.

In other words,

$$\frac{\psi(tA)}{\psi(tB)} = \frac{\psi(A)}{\psi(B)}$$

whenever this is rational. Going to a limit we obtain the same equality for arbitrary ratio, i.e. for arbitrary A, B and t. Hence

$$\frac{\psi(At)}{\psi(A)} = \frac{\psi(Bt)}{\psi(B)}$$

does not depend on A and B, and is a function only of t: $\psi(At) =$

$$\frac{\psi(At)}{\psi(A)} = \frac{\psi(Bt)}{\psi(B)}$$

$\psi(A)c(t)$ hence $\psi(t) = C_o t^p$ for some C_o and p. Therefore

$$a * b = \psi^{-1}(\psi a + \psi b) = (a^p + b^p)^{1/p}$$

for some p.

For scale-invariant functions the proof is essentially the same. QED.

Proof of Proposition 2

Let's consider first shift-invariant case. We demand that in case we combine Ae^{-kx} and $B\exp(-kx)$, we obtain $C\exp(-lx)$ for some C and l. Denoting $t = \exp(-kx)$ (as in proposition 1), we obtain that $At * Bt = (A * B)t^{p_2}$(where $p_2 = l/k$), for all A, B, t. In terms of ψ: if $\psi(A) + \psi(B) = \psi(C)$ then $\psi(tA) + \psi(tB) = \psi(t^{p_2}C)$. If $p_2 = 1$ we obtain the case of proposition 2, where $f * g$ is always of the same symmetry as f and g; but we consider only the case where there can be another symmetry, so $p_2 \neq 1$. Likewise, considering combination of three terms we conclude that if $\psi A + \psi B + \psi D = \psi E$ then $\psi(tA) + \psi(tB) + \psi(tD) = \psi(t^{p_3}E)$. So if $\psi A + \psi A = \psi C$ then due to $\psi(tA) + \psi(tB) = \psi(t^{p_2}C)$ we obtain that

$$\psi(t^{p_2}C) + \psi(tD) = \psi(t^{p_3}E).$$

On the other hand, since $\psi C + \psi D = \psi E$ we obtain that $\psi(tC) + \psi(tD) = \psi(t^{p_2}E)$.

Calculating the difference between these two equations, we conclude that

$$\psi(t^{P_2}C) - -\psi(tC) = \psi(t^{p_3}E) - -\psi(t^{p_2}E).$$

For arbitrary C, E we can find such A, B, D, so this is true for arbitrary C, E, t.

The right-hand side does not depend on C, therefore the left-hand side also does not depend on C and is a function of t only: $\psi(t^{p_2}C) - \psi(tC) = z(t)$.

Denoting $\widetilde{C} = tC$ and $\widetilde{t} = t^{p_2^{-1}}$, we obtain that for all $\widetilde{t}$, $\widetilde{C}$:

$$\psi(\widetilde{t}\widetilde{C}) - -\psi(\widetilde{C}) = \widetilde{z}(\widetilde{t}),$$

where $\widetilde{z}(\widetilde{t}) \stackrel{\text{def}}{=} z(t)$ and $t = (\widetilde{t})^{1/(p_2^{-1})}$. Hence $\psi(\widetilde{t}\widetilde{C}) = \psi(\widetilde{C}) + \widetilde{z}(\widetilde{t})$ and $\psi(\widetilde{t}\widetilde{C}) = \psi(\widetilde{C}\widetilde{t})$ implies that $\psi(\widetilde{C}) = \widetilde{z}(\widetilde{t}) = \psi\widetilde{t} + \widetilde{z}(\widetilde{C})$, hence $\psi(\widetilde{C}) - -\widetilde{z}(\widetilde{C}) = \psi(\widetilde{t}) - -\widetilde{z}(\widetilde{t})$ so $\psi - -\widetilde{z} =$ const, and $\widetilde{z}(\widetilde{t}\widetilde{c}) = \widetilde{z}(\widetilde{t}) + \widetilde{z}(\widetilde{C})$. Hence $\widetilde{z}(t) = ln\, t$ and $\psi(t) = ln\, t + C$, and $a * b = ab$, i.e. $* =$ product.

For scale-invariance the proof is similar. QED.

Proof of Propositions 3 and 4

Just like in proposition 1 from the demand that $\phi(A\exp(-kx)) = B\exp(-kx)$ we conclude that $\phi(At) = Bt$ i.e. $\phi(tA) = t\phi(A)$, hence for $A = 1$: $\phi(t) = \text{const} \times t$. Likewise $\phi(A\exp(-kx)) = B\exp(-lx)$ implies that $\phi(t) = \text{const} \cdot t^p$. QED

Proof of the Theorem

For Planck formula

$$P(x) = \frac{x^3}{e^{x/kT} - 1} :$$

add $(P) = 1$ (one addition), mult $(P) = 1$ (one division) and const $(P) = 1$. Let's prove that whatever other expression Q with this asymptotics we take – then always $P \preceq Q$ i.e. add $Q \geq 1$, mult $Q \geq 1$ and either add $Q > 1$ or mult $Q > 1$, or the third case of $>$ occurs.

1.° First let's prove that add $Q \geq 1$, i.e. add Q is not 0, i.e. it is impossible to construct the expression Q with given asymptotics by pure multiplication or : . Indeed, multiplication of Ax^p -type expression is again Ax^a; multiplication or division of shift-invariant expression leads also to $A\exp(-kx)$. So arbitrary expression with add $Q = 0$ is $Ax^a\exp(-kx)$. Asymptotic equivalence for $x \to \infty$ implies that $a = 3$, $k = 1/kt$, but then asymptotic in $x \to 0$ is wrong. So add $Q = 0$ is impossible hence add $Q \geq 1$.

2.° Prove now that mult $Q \geq 1$. Indeed, if mult $Q = O$ then we should have $Q = \Sigma A_i x^{a_i} + \Sigma B_i \exp(-k_i x)$ ($-$ can be changed into $+$ by changing signs of corresponding A_i or B_i). This expression cannot give correct asymptotics for $x \to \infty$, so mult $Q \geq 1$.

Let's now prove that if mult $Q = 1$ and add $Q = 1$ then Q has the same structure as P.

3.° mult $(P) = 1$ can mean either multiplication or division. Prove that it cannot be multiplication.

Indeed, in this case $Q = e_1 e_2 + e_3$ or $Q = e_1(e_2 + e_3)$ for some basic functions e_i. In the first case, if e_1, e_2 had symmetries of not different types then their product would have the same type of symmetry, so $e_1 e_2$ would be basic, hence we could obtain an expression with mult $Q = 0$ - and we proved that it is impossible. So they are of different types hence $e_1 e_2 = Ax^a\exp(-bx)$.

Here e_3 is either Cx^c, or $C\exp(-cx)$. In the first case $Q = Ax^a\exp(-bx) + Cx^c$; so asymptotics for ∞ implies that

$$\lim_{x\to\infty}[Ax^{a-3}\exp\left((\frac{1}{kT} - -b)x\right) + Cx^{c-3}\exp(x/kT)] = 1.$$

The second term tends to ∞. If $b > 0$ then the first term is asymptotically smaller so this sum $\to\infty$. If $b < 0$ then, vice versa, the second term can be neglected and the result is $\pm\infty$ (dependent on the sign of A).

In the second case $Q = Ax^a\exp(-bx) + C\exp(-cx)$ and asymptotics for ∞ implies that $a = 3, b = 1/kt$, but then at 0 asymptotics is wrong.

Likewise, simply analyzing all possible cases, we can prove that $Q = e_1(e_2 + e_3)$ is also impossible.

So Q must contain division.

4.° There are three ways to combine division and addition: $Q = e_1 + e_2/e_3$, $Q = (e_1 + e_2)/e_3$ and $Q = e_1/(e_2 + e_3)$. Due to the fact that for arbitrary basic function $e(x)$ its inverse $e^{-1}(x) = 1/e(x)$ is also basic, the first two expressions are of the type $Q = e_1 + e_2\tilde{e}_3$ and $Q = (e_1 + e_2)\tilde{e}_3$, where $\tilde{e}_3 = e_3^{-1}(x)$, i.e. of the already rejected types. So the only possibility remains – when $Q = e_1/(e_2 + e_3)$ i.e. Q is of the same structure as P.

5.° Now dependent on which of e_i are exponents and which are not we obtain the following cases:

a. all e_i are exponents. Then, dividing all parts by e_1, we obtain that

$$Q = \frac{1}{\tilde{e}_2 + \tilde{e}_3},$$

where $\tilde{e}_2, \tilde{e}_3$ are again exponents. For $x \to 0$ we get either const or $1/cx$ and not x^a.

b. $e_1, e_2 - -\exp, e_3-$ not. Then

$$\frac{Q = A\exp(\alpha x)}{B\exp(bx) + Cx^c.}$$

When $x \to 0$, this tends to A/B if $c > 0$ and to

$$\frac{A}{Cx^c} = \frac{A}{C}x^{-c} \quad \text{if } c < 0.$$

It should be x^2, so $c = - - 2$. For $x \to \infty$, if $b > 0$, then the asymptotics is purely exponential; if $b < 0$ then for $x \to \infty$: $B\exp(bx) << Cx^{-2}$, so $Q \sim \text{const} \cdot x^2\exp(\alpha x)-$ the wrong asymptotics. So this case also cannot occur.

c. $e_1 - -\exp, e_2, e_3$ – not. Then

$$Q = \frac{A\exp(ax)}{Bx^b + Cx^c}.$$

Without losing generality we can assume that $b > c$. Then for $x \to \infty : Bx^b >> Cx^c$, hence asymptotically $Q \sim (A/B)\exp(ax)x^{-b}$, so $b = -3$, $\alpha = 1/(kT)$. But when $x \to 0$, $Bx^b << Cx^c$, so asymptotic is $\sim A/C\, x^{-c}$, hence $c = -2$, and this contradicts to our assumption that $b > c$. So this case is also impossible.

We have enumerated all cases when $e_1 = \exp$, so e_1 is a degree. Likewise to a. – all three degrees are impossible, therefore one of e_1 is an exponent. In all these cases $P \preceq Q$. QED.