



Eugene Y. S. Chua*

Degeneration and Entropy

<https://doi.org/10.1515/krt-2021-0032>

Published online January 10, 2022

Abstract: Lakatos's analysis of progress and degeneration in the *Methodology of Scientific Research Programmes* is well-known. Less known, however, are his thoughts on degeneration in *Proofs and Refutations*. I propose and motivate two new criteria for degeneration based on the discussion in *Proofs and Refutations – superfluity* and *authoritarianism*. I show how these criteria augment the account in *Methodology of Scientific Research Programmes*, providing a generalized Lakatosian account of progress and degeneration. I then apply this generalized account to a key transition point in the history of entropy – the transition to an information-theoretic interpretation of entropy – by assessing Jaynes's 1957 paper on information theory and statistical mechanics.

Keywords: Lakatos, philosophy of physics, degeneration, entropy, information theory, Jaynes

As a young man I tried to read thermodynamics, but I always came up against entropy as a brick wall that stopped my further progress. I found the ordinary mathematical explanation, of course, but no sort of physical idea underlying it. No author seemed even to try to give any physical idea. Having in those days great respect for textbooks, I concluded that the physical meaning must be so obvious that it needs no explanation, and that I was especially stupid on the particular subject.
– James Swinburne (1904, p. 3)

My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name. In the second place, and more importantly, no one knows

*Corresponding author: Eugene Y. S. Chua, University of California San Diego, La Jolla, USA,
E-mail: eychua@ucsd.edu

Open Access. © 2021 Eugene Y. S. Chua, published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

what entropy really is, so in a debate you will always
have the advantage.

– Claude Shannon, according to McIrvine and Tribus (1971, p. 180)

1 Introduction

Lakatos (1976/2015) argued in *Proofs and Refutations* (P&R) that the comprehension of mathematical concepts must be accompanied with a clear understanding of how and why these concepts came into existence. For Lakatos, a concept is to be understood in terms of a temporally extended *process* through which the initial, primitive, concept is continually refined. To rip a concept apart from its context of discovery or its *problem-situation* – the problems or questions which led to the concept’s genesis and evolution – is to miss a complete understanding of it.

All of the above – concerning how to *comprehend* a concept – has been much discussed over the last few decades.¹ Less discussed is how to *evaluate* a concept according to the heuristic approach presented in P&R: how do we know whether a concept is problematic, needs rehabilitation, or, worse still, must be abandoned, given the heuristic approach? In short, how do we know whether a concept is *degenerating*?

That not much has been said about this is curious, since Lakatos clearly had some such standard in mind. In this paper, I explore Lakatos’ views on degeneration in P&R, which has often been neglected for the sort of degeneration Lakatos (1978) discussed in *Methodology of Scientific Research Programmes* (MSRP). It seems to me that P&R offers new criteria for degeneration that sheds light on Lakatos’s approach.

The primary goal here is to motivate an account of degeneration based on my reading of P&R. I propose two criteria for degeneration: *superfluity*, or generalization for generalization’s sake, which involves the introduction of trivial extensions or terminology to a theory or concept; and *authoritarianism*, the introduction and employment of a concept into the discussion without justification while ignoring the problem-situation of said concept. In my view, these notions of degeneration depart from the two conditions found in MSRP: the former relate not to the *content* of a concept or theory,² but to their *methodological* aspects (i.e. *depth*). As such, by considering the notions of

¹ See e.g. Corfield (1997), Leng (2002), and Werndl (2009).

² It seems to me that Lakatos did not really have a clear-cut distinction between concept and theory. For instance, he talks about the “theory of polyhedra” throughout P&R, but also refers to the “concept of polyhedron” as well. I see no cost to this lack of distinction here. Hence, I follow Lakatos in using “concept” and “theory” interchangeably in this paper whenever the context is clear. Thanks to Craig Callender and Kerry McKenzie for independently raising this concern.

degeneration in both P&R and MSRP, I propose an extended account of Lakatosian degeneration which evaluates both content and depth.

The secondary goal is to apply this extended account of degeneration to entropy. Why entropy? The concept of entropy has a tumultuous past, with many interpretations,³ complications,⁴ and disagreements colouring its rich history. This is coupled, however, with its extensive usage in countless sciences, be it in black hole thermodynamics,⁵ quantum theory,⁶ AdS/CFT research in theoretical physics,⁷ and even biology and neuroscience.⁸ There remains much room to evaluate these applications and extensions of the concept of entropy, including whether we *should* extend it in these ways.⁹ This makes entropy an interesting case study for degeneration – after all, my proposed account of degeneration is intended to evaluate the extension of a concept. Since the assessment of the concept of entropy remains pretty much an open question, it is hoped that my account will provide some heuristics for beginning this assessment. Here, as proof of concept, I focus on one key transition point for the concept of entropy – the transition from thermodynamic to information-theoretic interpretations of entropy. In Lakatosian style, I identify a key piece of writing in this transition and evaluate it with the twin criteria developed here. By critiquing Jaynes’s (1957) landmark paper on thermodynamics and information, I argue that this transition suffered from superfluity and authoritarianism, and hence, degeneration.

2 Degeneration in P&R

P&R takes the form of a dialogue between a fictional teacher and his students which includes Gamma (and later Alpha, Rho and Zeta, among others). Lakatos’s main target was a ‘deductivist’ approach to mathematics:

This style starts with a painstakingly stated list of axioms, lemmas and/or definitions. The axioms and definitions frequently look artificial and mystifyingly complicated. One is never told how these complications arose. The list of axioms and definitions is followed by the carefully worded theorems. These are loaded with heavy-going conditions; it seems

3 See Uffink (2001) for how interpreting the original thermodynamic concept of entropy is itself a challenge.

4 See e.g. Callender (1999), Chua (Forthcoming), or Goldstein et al. (2020) for some of these complications concerning how different entropies relate to others.

5 See e.g. Dougherty and Callender (1319) or Wallace (1529).

6 See e.g. Bub (2005).

7 See e.g. Natsuume (2015).

8 See Friston and Stephan (2007).

9 See e.g. Roach (2020) who evaluates the use of entropy in biology.

impossible that anyone should ever have guessed them. The theorem is followed by the proof. (Lakatos 1976/2015, p. 151)

In Lakatos's view, this approach to mathematics is misguided. By rationally reconstructing the historical development of the Euler characteristic:¹⁰

$$V - E + F = 2$$

he showed that these definitions, axioms, theorems and so on developed only as the result of a long history of proofs and refutations: they are *proof-generated concepts*.

For Lakatos, actual mathematics is not deductivist. Nevertheless, it is a *rational* affair representative of what he called the *heuristic approach*:

...deductivist style tears the proof-generated definitions off their 'proof-ancestors', presents them out of the blue, in an artificial and authoritarian way. It hides the global counterexamples which led to their discovery. Heuristic style on the contrary highlights these factors. It emphasises the *problem-situation*: it emphasises the 'logic' which gave birth to the new concept. (Lakatos 1976/2015, p. 153)

More generally, a concept is only appropriately understood when we understand its historical trajectory. In Lakatos's view, "there is a simple pattern of mathematical discovery – or of the growth of informal mathematical theories" (Lakatos 1976/2015, pp. 135–136), which is given by the following seven stages:

- (1) *Primitive conjecture.*
- (2) *Proof* (a rough thought-experiment or argument, decomposing the primitive conjecture into subconjectures or lemmas).
- (3) "*Global*" *counterexamples* (counterexamples to the primitive conjecture) emerge.
- (4) *Proof re-examined.*
- (5) *Proofs of other theorems are examined to see if the newly found lemma or the new proof-generated concept occurs in them.*
- (6) *The hitherto accepted consequences of the original and now refuted conjecture are checked.*
- (7) *Counterexamples are turned into new examples – new fields of inquiry open up.*

Call this the *heuristic process*. This tells us how a mathematical theory or concept ought to *grow*, from a rough primitive conjecture to a proof-generated concept (and beyond), through the use of heuristics like employing counterexamples, discovering hidden lemmas, and so on. We appropriately comprehend a concept only

¹⁰ The Euler characteristic was initially postulated to universally describe polyhedra in terms of their number of faces (F), vertices (V), and edges (E).

when we start with the primitive conjecture – the genesis of the concept – and grasp the ensuing adjustments and responses to the concept through which it is precisified and stretched.

But as Gamma points out, *growth* is opposed to *degeneration* (Lakatos 1976/2015, p. 103). However, while Lakatos paints a clear picture as to how mathematical theories grow, he is much less explicit about degeneration. This is curious because the term “degeneration” seems to bear significant normative weight in Lakatos’s appraisal of research methodologies. My goal here is to remedy this situation by explicating Lakatos’s account of degeneration in terms of two distinct criteria.

2.1 Superfluity

The first criterion for degeneration appears when Alpha charts the development of the dialogue in P&R thus far, in response to ever-exotic counter examples (Lakatos 1976/2015, p. 86):

- (1) one vertex is one vertex.
- (2) $V = E$ for all perfect polygons.
- (3) $V - E + F = 1$ for all normal open polygonal systems.
- (4) $V - E + F = 2$ for all normal closed polygonal systems, i.e. polyhedra.
- (5) $V - E + F = 2 - 2(n - 1)$ for normal n -spheroid polyhedra.
- (6) $V - E + F = 2 - 2(n - 1) + \sum_{k=1}^F e_k$ for normal- n -spheroid polyhedra with multiply-connected faces.
- (7) $V - E + F = \sum_{k=1}^F \{2 - 2(n_j - 1) + \sum_{k=1}^F e_{kj}\}$ for normal n -spheroid polyhedra with multiply-connected faces and with cavities.

Alpha proclaims that “this is a miraculous unfolding of the hidden riches of the trivial starting point [(1)]”. However, Rho retorts: “Hidden ‘riches’? The last two only show how cheap generalisations may become!” Gamma concurs:

(6) and (7) are not growth, but *degeneration!* Instead of going on to (6) and (7), I would rather find and explain some exciting new counterexample [to $V - E + F = 2$!] (Lakatos 1976/2015, p. 103, emphasis mine)

Degeneration is tied, in part, to “cheap generalisation” in the process of developing a concept or theory. It plays an evaluative role insofar as it tells us when, in something like the chain of generalizations from (1) to (7), we should stop in response to a new counterexample. Gamma describes (7) as a pointless generalization: who cares about polyhedrons with cavities or multiply-connected faces?

To Gamma, “It serves only for making up complicated, pretentious formulas for nothing (Lakatos 1976/2015, pp. 102–103).”

A cheap generalization, in Gamma’s view, is a *trivial extension* of a concept (Lakatos 1976/2015, p. 102). The generalization was not well-motivated, and thus nothing *deep* was gleaned from doing so, even though our formula does become more general. It is, to put it bluntly, *generalization for the sake of generalization*. When a concept is extended in order to encompass new cases, but those cases were not relevant to the original problem-situation (or the heuristics that followed), Gamma would consider them trivial extensions.

Lakatos, channelling Pólya, assents to this in his description of what cheap generalization amounts to:

Pólya points out that shallow, cheap, generalisation is ‘more fashionable nowadays than it was formerly’. It dilutes a little idea with a big terminology. The author usually prefers to take even that little idea from somebody else, refrains from adding any original observation, and avoids solving any problem except a few problems arising from the difficulties of his own terminology.

Another of the greatest mathematicians of our century, John von Neumann, also warned against this ‘danger of degeneration’, but thought it would not be so bad ‘if the discipline is under the influence of men with an exceptionally well-developed taste’. One wonders, though, whether the ‘influence of men with an exceptionally well-developed taste’ will be enough to save mathematics in our ‘publish or perish’ age. (Lakatos 1976/2015, p. 104, fn. 160, emphasis mine)

For our purposes, it certainly helps that Lakatos directly connects “shallow, cheap, generalisation” to the “danger of degeneration”. Alternatively stated, the generalization of a concept to new domains without justification is *superfluous* and only adds unnecessary terminology (“pretentious formulas”). This, in Lakatosian terms, is concept-stretching “gone wrong” – a concept is stretched too far without justification, resulting in trivial generalizations. This leads to the concept’s degeneration. Call this first criterion of degeneration *superfluity*.

What counts as justification and can avoid superfluity? Of course, the justification must be relevant to the problem-situation at hand,¹¹ but what else? We can distil further insights from Lakatos’s discussion of the Euler characteristic: in that case, the problem-situation is one where the key concern is about classifying (what we intuitively might call) *polyhedra*. However, while cavities were helpful in categorizing polyhedra based on *Definition 1*, i.e. the *naïve* notion of a polyhedron as a solid, a single “polyhedron with a cavity” corresponds to an entire class of

¹¹ Some might argue that this is too restrictive – what if a proposal has relevance elsewhere? Then it must prove its relevance to *that* problem-situation.

polyhedra in the succeeding proof-generated concept of polyhedra as connected surfaces (Lakatos calls this *Definition 2*) (Lakatos 1976/2015, p. 16). On this definition, a polyhedron is not a *solid*. Given this background, the number of cavities does not actually pick out a unique type of polyhedra.¹² Hence any extension of the Euler characteristic which takes into account the number of cavities simply plays no real role in advancing the research of polyhedra.¹³ Instead, we have merely added superfluous terminology which does not concern the objects of interest.

From this we might infer that the sort of justification which vindicates any particular generalization or introduction of terminology is one that can motivate why this terminology or generalization was introduced – particularly in relation to the sort of objects or concepts we care about – and how it can possibly lead to the growth of the theory or concept. It cannot be just a *trivial pun*, which is essentially what a “polyhedron with a cavity” is, since it is not a polyhedron *per se* at all given the problem-situation where *Definition 2* is accepted as a proof-generated concept – it must serve some *use* and justify its own existence, so to speak.

Of course, this requirement is not precise, for it is not always clear what is trivial. Alpha questions: “You may be right after all. But who decides where to stop? Depth is only a matter of taste (Lakatos 1976/2015, p. 103).” In response, Gamma proposes:

Why not have mathematical critics just as you have literary critics, to develop mathematical taste by public criticism? We may even stem the tide of pretentious trivialities in mathematical literature. (Lakatos 1976/2015, p. 104)

Unfortunately, Lakatos does not say much more about what “taste” amounts to, though his comment on von Neumann’s comment suggests skepticism towards its role.

Nevertheless, hindsight helps, as in the case of cavities failing to pick out the relevant sort of properties for polyhedra: “relevance” is determined as a matter of practice, adoption, and actual contribution to the problem-situation, e.g. *Definition 2* as an improvement over *Definition 1*. As Kiss puts it succinctly (though in the context of the MSRP): “One step in a research program can be treated as progressive or degenerating only in hindsight, when we see future developments. Appraisal of research programs is as fallible as the theories themselves (Kiss 2006, p. 316).”

¹² See Lakatos (1976/2015), p. 97, fn. 150(a).

¹³ Likewise for multiply-connected faces: it was simply not interesting for research because our means of classifying polyhedra simply did not require multiply-connected faces. As Lakatos notes, as the theory of polyhedra evolved, “the problem of how multiply-connected faces influence the Euler-characteristic of a polyhedron lost all its interest” (Lakatos 1976/2015, p. 97, fn. 150(c)).

We can do better in other cases. In line with the heuristic approach, conscious awareness of the problem-situation – and the heuristics associated with it – allows one to avoid superfluity. Lakatos discusses the example of the mathematician Becker, who aimed to provide a conclusion to the classification problem by providing a new generalization to the Euler characteristic:

$$V - E + F = 4 - 2n + q$$

where n is the number of cuts that is needed to divide the polyhedral surface into simply-connected surfaces for which $V - E + F = 1$, and q is the number of diagonals that one has to add to reduce all faces to simply-connected ones (Lakatos 1976/2015, p. 103, fn. 158). Unfortunately for Becker, Lakatos notes that the mathematicians Lhuillier and Jordan had already written about this over half a century ago, except in different terminologies. While Becker's work does count as a valid generalization of Euler's original formulation, it was ultimately trivial – adding only new terminology – and did not contribute to the development of the concept. Here it is clear that a cognizance of the problem-situation would have helped: by being aware of the concept's past iterations, problems, and errors, one learns what *not* to do.

By being aware of the problem-situation for the concept, we can avoid trivial generalizations. Beyond that, we can only hope to pinpoint degeneration in retrospect unless we are blessed with the great gift of “exceptional taste” in the subject matter at hand (in which case, a certain sort of clairvoyance – of what *will* work out in research – is possible). However, given Lakatos's distaste for formalism and inclination towards informal heuristics – which are themselves not always precise – this might be to Lakatos's liking.

2.2 Authoritarianism

In discussing the heuristic approach, Lakatos notes a common response to byzantine definitions (in this case, Carathéodory's definition of a measurable set) with disapproval:

Of course there is always the easy way out: mathematicians define their concepts just as they like. But *serious teachers do not take this easy refuge*. Nor can they just say that this is the correct, true definition ... and that mature mathematical insight should see it as such. (Lakatos 1976/2015, p. 162)

There is something problematic with the deductivist style of simply introducing concepts out of thin air without appropriately situating them within the proof's problem-background. This approach is *authoritarian*:

One can easily give more examples, where stating the primitive conjecture, showing the proof, the counterexamples, and following the heuristic order up to the theorem and to the proof-generated definition would dispel the *authoritarian mysticism of abstract mathematics*, and *would act as a brake on degeneration*. A couple of case-studies in this degeneration would do much good for mathematics. Unfortunately the deductivist style and the atomisation of mathematical knowledge protect ‘*degenerate*’ papers to a very considerable degree. (Lakatos 1976/2015, p. 163, emphasis mine)

This provides a new criterion for degeneration, which occurs when a concept (or terminology, or definition, or perhaps even a theory) is introduced without justification into some line of inquiry. New concepts are instead used with the attitude “that this is the correct, true definition” without qualification. This adds a “mystical” and “authoritarian” element to these new concepts which ignores the background problem-situation leading to that line of inquiry to begin with. Call this criterion for degeneration *authoritarianism*.

While many of Lakatos’s examples of authoritarian methodology were textbooks (like Rudin or Halmos), I should emphasize that there is no reason to interpret his discussions about authoritarianism as something which *only* applied to pedagogical works.¹⁴ For instance, in the quote above, Lakatos was clearly discussing how the dispelling of “authoritarian mysticism” is needed to remove the protection of degenerate *papers*, i.e. research, by the deductivist style of mathematics. Furthermore, in Appendix I, Lakatos writes that:

It was the infallibilist philosophical background of Euclidean method that bred the authoritarian traditional patterns in *mathematics*, that prevented publication and discussion of conjectures, that made impossible the rise of mathematical criticism. (Lakatos 1976/2015, p. 147, emphasis mine)

Lakatos is here using authoritarianism as a criterion for evaluating *mathematics as a whole*, including research like publications and discussions, rather than pedagogy in particular. As such, I believe that authoritarianism should be interpreted as a criterion of evaluating degeneration which is applicable to research as well as pedagogy.

Lakatos raises the example of Rudin’s discussion of bounded variation. While introducing the Riemann-Stieltjes integral, Rudin introduces the notion of bounded variation.¹⁵ He then proves a theorem to the effect that a function of bounded variation, satisfying other criteria, is also a member of the class of Riemann-Stieltjes integrable functions. However, Lakatos accuses Rudin of failing to explain why the Riemann-Stieltjes integral and bounded variation were relevant to begin with:

¹⁴ I thank an anonymous reviewer for pushing me to clarify this.

¹⁵ For Lakatos’s discussion, see Lakatos (1976/2015), pp. 155–162.

So now we have got a theorem in which two mystical concepts, bounded variation and Riemann-integrability, occur. But two mysteries do not add up to understanding. Or perhaps they do for those who have the “ability and inclination to pursue an abstract train of thought”? (Lakatos 1976/2015, p. 156)

The *non-degenerative* way of presenting the concept would have shown that the two concepts arose as proof-generated concepts out of the same problem-situation:

A heuristic presentation would show that both concepts – Riemann-Stieltjes integrability and bounded variation – are proof-generated concepts, originating in one and the same proof: Dirichlet’s proof of the Fourier conjecture. This proof gives the problem-background of both concepts. (Lakatos 1976/2015, p. 156)

Lakatos notes that Rudin *does* mention this, but in a way that is disconnected from the two aforementioned concepts: it was hidden in some exercise in a different chapter. Lakatos declares that the two concepts, introduced this way, were introduced in “an authoritarian way (Lakatos 1976/2015, p. 156, fn. 12)”.¹⁶

Thus, both superfluity and authoritarianism arise from a failure to grapple with the problem-situation. While superfluity arises from trivial generalizations of concepts arising from this lack of awareness, authoritarianism arises from the unmotivated introduction and application of concepts.

For Lakatos, to “introduce” a term out of the blue into a discussion is a “magical operation which is resorted to very often in history written in deductivist style! (Lakatos 1976/2015, 157, fn. 17)”. To treat bounded variation or the Riemann-Stieltjes integral as an “introduced” concept rather than a proof-generated one – as he shows in Lakatos (1976/2015, pp. 156–162) – is to miss out on the understanding of the concept. On the heuristic approach,

[...] the two mysterious definitions of bounded variation and of the Riemann-integral are *entzaubert*, deprived of their authoritarian magic; their origin can be traced to some clear-cut problem situation and to the criticism of previous attempted solutions of these problems. (Lakatos 1976/2015, p. 158)

In sum, I understand authoritarianism – the second criterion of degeneration – as introducing new concepts into some line of inquiry without justification, while

16 Some might worry that having a complete grasp or presentation on any problem-situation might be too demanding. However, Lakatos’s reply to this “pedestrian argument” is: “let us try (Lakatos 1976/2015, p. 153).” He thinks of this stringency as a positive point: while this makes academic work demanding and long, “. . . it has to be admitted that they would be much fewer too, as the statement of the problem-situation would too obviously display the pointlessness of quite a few of them (Lakatos 1976/2015, p. 153, fn. 5).” Thanks to Nuhu Osman Attah for raising the worry.

ignoring the problem-situation and the heuristics that led us to that discourse to begin with.

2.3 Their Normative Import

Recall that the heuristic approach places emphasis on understanding the background and historical trajectory of a concept, over and above the concept in its current form. The original problem and the errors that followed (i.e. the problem-situation) are just as important as the end-product and the proof-generated concept, because they tell us what has not worked (and will not work), why the concept is the way it is now, and hopefully the available routes for development based on those errors (at least, it rules out the *unavailable* routes for development).

By failing to grasp this problem-situation, superfluity and authoritarianism miss out on a complete understanding of the concept as *part of a historical trajectory*, instead atomizing it as a stand-alone concept – methodologies with superfluity and authoritarian tendencies thus hinder an understanding of the concept. As Zeta puts it, “A problem never comes out of the blue. It is always related to our background knowledge (Lakatos 1976/2015, p. 74)”. By adopting a methodology which elides this background knowledge, our understanding of a concept and the associated problem is rendered incomplete.

On one hand, authoritarianism fails to account for past problems and errors by tearing apart present discussions from past problems – the discussion is presented without context; the reader is told to accept it on faith, or that they need “mathematical maturity” to understand it (Lakatos 1976/2015, p. 151, fn. 1). This obscures the errors and problems crucial to generating the proof-generated concept. We then lose sight of the direction the concept was taking in the long chain of proofs and refutations: we are “rewriting history to purge it from error” (Lakatos 1976/2015, p. 49) and hence “the zig-zag of discovery cannot be discerned in the end-product” (Lakatos 1976/2015, p. 44). The degenerate concept becomes atomized as a result, which hinders growth in Lakatos’s view.

On the other hand, superfluity reflects a lack of concern for a concept’s problem-situation. Terminology is produced, but not because it presents an insightful development for the concept and its trajectory. Sometimes, as in Becker’s case, one mistakes themselves to be presenting fruitful development for a concept. Again, this approach treats the concept at hand as one that is divorced from its problem-situation – instead of considering what problems the terminology is meant to resolve, we are instead pursuing “cheap, shallow generalisations” even if the result is ultimately trivial. We have seen this in the case of generalizing the Euler characteristic to (6) and (7): obviously, we *can* generalize the Euler

characteristic to the case of intuitive polyhedra with cavities and multiply-connected faces, but the naïve terminology – of cavities and multiply-connected faces – and the accompanying generalizations were simply no longer fruitful to the discussion of polyhedra at that point of the trajectory of the concept in the dialogue. A superfluous extension of a concept involves ever more esoteric “generalizations” to cover cases no one cares about (in the context of the line of inquiry surrounding that concept). In Lakatos’s words:

Quite a few mathematicians cannot distinguish the trivial from the non-trivial. This is especially awkward when a lack of feeling for relevance is coupled with the illusion that one can construct a perfectly complete formula that covers all conceivable cases. Such mathematicians may work for years on the ‘ultimate’ generalisation of a formula, and end up by extending it with a few trivial corrections. (Lakatos 1976/2015, p. 103, fn. 158)

By failing to grasp what is trivial (which can be aided by hindsight or a grasp of the problem-situation), research degenerates by either treading trodden grounds (as with the case of Becker) or extending a concept to domains which are simply unfruitful (as in the case of cavities and multiply-connected faces).

In short, both superfluity and authoritarianism have clear normative import: if our goal is to pursue growth for the concept by having a clearer understanding of the concept, we ought to avoid both forms of degeneration. Degeneration in these two senses thus play an *evaluative* role for the growth of a concept.

3 Degeneration in P&R versus Degeneration in MSRP

I have focused on degeneration in the context of P&R, but how does this connect with Lakatos’s much more famous classification of scientific research programmes as degenerative or progressive? For Lakatos in MSRP, scientific theories should be understood akin to the heuristic approach to mathematical theories and concepts. They are not isolated atoms but *sequences* of theories or concepts – what he later calls a *scientific research programme* – grouped together by various criteria (such as their positive and negative heuristics).¹⁷ This “shifts the problem of how to appraise *theories* to the problem of how to appraise *series of theories*. Not an isolated *theory*, but only a series of theories can be said to be scientific or unscientific: to apply the term “scientific” to one *single* theory is a category mistake (Kiss 2006, p. 34).”

¹⁷ See Lakatos (1978), pp. 47–52.

The trajectory of this sequence of theories over time (i.e. its “problemshift” from one theory to another) is progressive or degenerative according to two criteria: whether it is (i) *theoretically* progressive and (ii) *empirically* progressive (Lakatos 1978, pp. 33–34). Being theoretically progressive refers to a succeeding theory containing “excess empirical content” by predicting novel facts compared to its predecessor, and would distinguish, according to Lakatos, a “scientific” problemshift from a non-“scientific” one. Being empirically progressive refers to the excess empirical content of this succeeding theory leading to the discovery of new facts, thereby corroborating the new theory’s novel predictions. A problemshift is deemed *overall progressive* if it is *both* theoretically and empirically progressive, and *overall degenerating* if it is not.

Much ink has been spilled over this point.¹⁸ What I am interested in is how the account of degeneration presented in MSRP can be augmented by the account of degeneration I have presented based on P&R, and how they can be collectively marshalled for the philosopher of science despite the obvious differences between the sciences and mathematics.¹⁹

There might be some doubt as to whether Lakatos’s views on mathematics in P&R can be so neatly transplanted to the scientific context. Despite the differences between science and mathematics, however, Lakatos emphasizes: “Mathematical heuristic is very like scientific heuristic – not because both are inductive, but because both are characterised by conjectures, proofs, and refutations. The important difference lies in the nature of the respective conjectures, proofs (or, in science, explanations), and counterexamples.” (Lakatos 1976/2015, p. 78) In my view, Lakatos is likely to take the heuristic approach to be applicable to both physics (and the sciences) and mathematics.

I believe the two accounts of degeneration complement each other. While the account in MSRP focused on *content*, the account in P&R focused on *depth* – how deep or trivial the research is, how it connects with its predecessors, the potential or actual fruitfulness of the research, and so on (discussed above in Section 2). Depth in turn hinges on *methodology*.

18 For an excellent collection of essays on Lakatos’s methodology, see Kamps, Kvasz and Stöltzner (2002).

19 Some might object that there are some differences between the sciences and mathematics that prevent such analysis. Interestingly, Hallert (1979a, 1979b), (Goldstein et al. 2020) and Stöltzner (2002) did something to similar effect but in the opposite direction of what I am doing here. They do the reverse, by proposing that we can apply the application of the conditions of progress and degeneration described in MSRP to mathematics (and mathematical physics) profitably. What I am doing here can be seen as a continuation of that sort of project – to bring Lakatos’s insights from his discussion of informal mathematics into physics. My analysis of Jaynes is an attempt to demonstrate that such an analysis is possible.

This former account of degeneration with respect to *content* takes a central role in Lakatos's project for scientific research programmes. Furthermore, it can be straightforwardly cashed out in terms of the *number* of theoretical predictions a theory makes and actually corroborated predictions relative to its rivals and competitors. These properties make them an easy target for analysis: for starters, we can just count the number of propositions (or sentences, or whatever your favourite truth-bearers are) non-vacuously entailed by the theory (and whether they are corroborated)!²⁰ This is not to say that considerations about content is somehow unimportant. We need yardsticks for discussing, comparing, and evaluating the content produced by scientific research programmes. If these yardsticks are clear, all the better. However, the focus on content – in terms of the predictions of a theory – does not really consider the *methodological* issues that might also be considered progressive or degenerating. I think this is important. In Rho's words in P&R: "Not every increase in content is also an increase in depth (Lakatos 1976/2015, p. 103)." A theory might have superior content while simultaneously experiencing a degeneration in methodology and *depth* of research.

This is where my proposed account of degeneration comes into play. This focuses on the *depth* of research at each problemshift – and this of course depends much more on the style and methodology of research, which may also be far more diverse than a generally accepted theory and its contents. Nevertheless, just as there are authoritative interpretations of theories even when there are generally numerous interpretations of any single theory, there are also authoritative figures, presentations, rhetoric, and methodologies, which may yet be open to analysis of depth. (An attempt to analyze one such authoritative presentation is made later in Section 4.) This, in turn, requires analysis of notions like taste, triviality, fruitfulness, awareness of the problem-situation and so on, as we have discussed in Section 2, which are not obviously amenable to logical analysis unlike content-oriented notions of degeneration.

But both *are* inseparable aspects of scientific research. We need to be concerned about both the depth of research, in terms of whether authoritarianism and superfluity are occurring, and the content being produced by the research, in terms of whether there is theoretical and empirical progress.

If I am right, we can extend Lakatos's classification of scientific research programmes quite straightforwardly: a problemshift is *overall progressive* if it is overall progressive with respect to content (content-degeneration) *and* avoids degeneration with respect to depth (depth-degeneration), that is, by being theoretically and empirically progressive *while* avoiding authoritarianism and superfluity. It is *degenerating otherwise*. Depth-degeneration comes in degrees – not all research at any one time will typically contain authoritarianism and superfluity,

²⁰ Of course, what counts as "non-vacuous" may yet be a matter for contention.

but *how much* research falls afoul of authoritarianism and superfluity will determine the degree of degeneration with respect to depth. My account of degeneration, based on the P&R, thus augments the account of degeneration found in MSRP and provides a new dimension of analysis for Lakatos's overall framework.

4 Entropy: Degeneration from Physics to Information

As a proof of concept, I begin this project by analyzing one important shift for the concept of entropy – the incorporation of information-theoretic notions into entropy by Jaynes (1957).²¹ Entropy plays a role of ever-increasing importance in our best sciences. Despite its ubiquity, the concept of entropy is not easily grasped. Some like Swinburne complained that “there is no sort of physical idea underlying [entropy]” (James 1904, p. 3); the concept of entropy does not come equipped with an obvious physical idea for us to latch onto, despite being defined in terms of physical quantities. This hints at degeneration – how can a concept that is so ubiquitous in physics be so imprecisely understood? In what follows, I examine the concept of entropy with Lakatos's method in the Appendix of P&R: highlight a piece of work which significantly influenced a concept's trajectory and point out its various degenerative traits, while suggesting what could have been done otherwise.

4.1 Jaynes's “Information Theory and Statistical Mechanics”

Jaynes was not the first to propose a marriage between information theory and statistical mechanics – that honour goes to Leon Brillouin.²² However, Jaynes's paper is one of the (if not *the*) most influential. As Seidenfeld (1986, p. 468) notes, “I doubt there is a more staunch defender of the generality of entropy as a basis for quantifying (probabilistic) uncertainty than the physicist E. T. Jaynes.” In a footnote to his famous 1973 paper on black hole entropy, Bekenstein observed that “the derivation of statistical mechanics from information theory was first carried out by E. T. Jaynes (Bekenstein 1973, fn. 17)”. In that same paper, he notes that by 1973, “The connection between entropy and information is well known (Bekenstein 1973,

²¹ I focus solely on the first of two papers he published on the topic in 1957, as the second part focuses largely on applying the account developed in part I, rather than presenting any novel arguments for the account itself.

²² See e.g. Brillouin (1956), Ch. 12 and beyond.

p. 2335)”, and later states, in a matter-of-fact way, that entropy “is the uncertainty in one’s knowledge of the internal configuration of the body (Bekenstein 1973, p. 2339).” Clearly, Jaynes’s information-theoretic subjectivist interpretation of statistical mechanics had won out by the 1970s – entropy has transformed from a quantity that keeps track of the reversibility or irreversibility of thermodynamic processes to a quantity that keeps track of the amount of information we have (or have lost) about said processes.

This makes Jaynes’s paper the perfect candidate for evaluating the degeneration in the transition from entropy as a thermodynamic, physical, concept about physical systems to an information-theoretic, subjective, concept about our ignorance or partial knowledge about physical systems.

Before Jaynes, the Gibbsian approach to statistical mechanics was by far the dominant paradigm in physics. Under the Gibbsian approach, the Gibbs entropy S_G of a physical system is defined by:

$$S_G = -k \int_{\Gamma} \rho(x, t) \log \rho(x, t) dx$$

Here, Γ is the $6N$ -dimensional phase space of the physical system in question, x is a point in Γ , dx is the volume element of Γ , and, importantly, $\rho(x, t)$ is some probability distribution defined over Γ which may or may not change over time. The Gibbs entropy is thus a function of these probability distributions. To define $\rho(x)$, consider a fictitious infinite ensemble of systems having (generally differing) microstates (position and momentum) consistent with the known macrostates (e.g. temperature, volume, pressure) of the actual system. In short, the macrostate(s), together with the dynamics of course, determine the choice of $\rho(x, t)$. S_G , in turn, is supposed to match the thermodynamic entropy at the thermodynamic limit,²³ which justifies its definition and also provides a physical basis for using it. Just as the (change in) thermodynamic entropy tracked the reversibility or irreversibility of thermodynamic processes, being equal to zero for reversible processes and greater than zero for irreversible ones (for closed systems over time), so does S_G at the appropriate limit.

Setting aside the debate over the nature of these fictitious ensembles (among other conceptual issues with the Gibbsian approach) for present purposes,²⁴ the approach so far is physical and world-oriented. In particular, the probability distributions $\rho(x)$ depend on the physical state of the system and are empirically determined – for instance, a system in equilibrium with an arbitrarily large heat bath is described with the canonical ensemble distribution, an isolated system

²³ This is when the number of particles in the system and the volume of the system itself approach infinity, with the ratio between particle number and volume held constant.

²⁴ See, again, Callender (1999) and Goldstein et al. (2020) for a nice overview of the issues with the Gibbsian approach.

with constant energy is described with the microcanonical ensemble distribution, and so on. These distributions then tell us the probability of some set of microstates obtaining given said constraints. We are here concerned simply with whether (and how likely) certain microstates of the actual physical system occur. Nothing in the Gibbsian theory *forces* us to employ notions of ignorance or knowledge so far, i.e. notions that would typically be described as “subjective” do not need to be employed in Gibbsian statistical mechanics.²⁵

In contrast, Jaynes (1957) explicitly introduces the notion of “subjective statistical mechanics” – where the usual rules of statistical mechanics can be “justified independently of any physical argument, and in particular independently of experimental verification (Jaynes 1957, p. 620)”. For Jaynes, statistical mechanics should not be interpreted as a physical theory in itself, with its equations, choice of distributions, and rules of computation justified by physical reasoning. Rather, it should be interpreted as a system of statistical inference, concerned primarily with our partial knowledge about physical systems. This system is then underpinned by the maximum entropy principle, which prescribes maximizing entropy as a formal means of representing maximal ignorance about that which we do not know. This principle is intended by Jaynes as an *a priori* principle of reasoning. The physics provides only the means of enumerating the possible states of the system and their properties (Jaynes 1957, p. 620). We use these as constraints on our knowledge (or lack thereof) and infer from these a set of equations via an appeal to information theory and subjectivist interpretations of probability and entropy.

Jaynes makes two related but distinct claims about his proposed account of statistical mechanics. First, statistical mechanics should be interpreted in a *subjectivist* fashion. This is opposed to an *objectivist* approach, which treats the probabilities produced by statistical mechanics as objective chances about events in the world (independent of what we think about those events). In his words:

The “subjective” school of thought regards probabilities as expressions of human ignorance; the probability of an event is merely a formal expression of our expectation that the event will or did occur, based on whatever information is available. (Jaynes 1957, p. 622)

For the subjectivist interpretation of statistical mechanics, the probability distributions, such as the canonical ensemble, the grand canonical ensemble, the microcanonical ensemble etc. are used to represent our partial knowledge of the

²⁵ There is a debate about whether the use of coarse-graining in Gibbsian statistical mechanics (which is necessary to recover the second law) might be anthropocentric and hence “subjective” – see Robertson (2020) for a discussion of why that is not necessarily the case (and hence why the Gibbsian approach still need not appeal to subjectivist notions).

system given certain constraints. The probabilities given by these ensembles are not really about the objective chances of the microstates occurring *per se*. Rather, these probabilities are interpreted to represent the degrees of belief we ought to have about these microstates, given suitable constraints.

What is important, however, is that the suitable constraints are not merely physical ones given by how the system is set up or behaving. This is his second major claim.

In addition to the subjectivist interpretation of statistical mechanics, Jaynes famously proposed an additional constraint to the inferential process: the maximum entropy principle (or MAXENT) (Jaynes 1957, p. 623). In short, the principle calls for the maximization of the entropy of the system, in addition to the other relevant physical constraints. However, the entropy of the system is interpreted in an information-theoretic way, *over and above* the subjectivist interpretation. Recall the Gibbs entropy:

$$S_G = -k \int_{\Gamma} \rho(x, t) \log \rho(x, t) dx$$

Under the subjectivist interpretation, $\rho(x, t)$ now represents the degrees of belief we ought to have in some (set of) microstates x obtaining at time t . Over and above that, S_G is to be interpreted (with the Boltzmann constant k set to unity via a choice of units) as the (continuous version of) Shannon entropy instead, representing “uncertainty” contained in $\rho(x, t)$, i.e. the “uncertainty” contained within our degrees of belief about said system. The intuition is that a peaked distribution contains less uncertainty than a flat distribution, and it turns out that S_G for a peaked distribution is indeed lower than a flat distribution (see Figure 1. for a visual aid). Furthermore, just as collecting information is additive, so too is S_G additive (as a simple result of its logarithmic form).

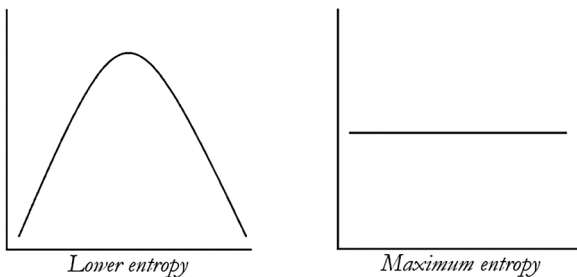


Figure 1: A schematic representation of a “peaked” distribution versus a “flat” one.

In sum, MAXENT is treated as an *a priori* rationality constraint: we assume maximal ignorance about a system except what we know about it, where “maximal ignorance” is equated to adopting a maximum information-entropy constraint on the system. This is what Jaynes meant by his account of statistical mechanics being a general account of statistical inference (Jaynes 1957, p. 621, pp. 629–630). Jaynes then showed that the adoption of MAXENT can recover all the usual equations and expressions of statistical mechanics.

As with Lakatos’s account in MSRP, much has already been said about the content of Jaynes’s claims, and I will not add more to the mix.²⁶ I focus on the methodological *depth* of Jaynes’s paper instead by applying the extended account of growth and degeneration.

I begin with the problem-situation. The founding motivations of statistical physics, found in the works of Boltzmann and Gibbs, are quite clear: to understand the molecular foundations of thermodynamics, and to interpret thermodynamics in terms of molecular mechanics.²⁷ As Boltzmann states in the introduction to his *Lectures on Gas Theory*:

I hope to prove in the following that the mechanical analogy between the facts on which the second law of thermodynamics is based, and the statistical laws of motion of gas molecules, is also more than a mere superficial resemblance. (Ludwig 1896/1995, p. 5)

Gibbs, while differing in his approach to statistical mechanics, held a similar view about the goal of statistical mechanics:

We may [...] confidently believe that nothing will more conduce to the clear apprehension of the relation of thermodynamics to rational mechanics, and to the interpretation of observed phenomena with reference to their evidence respecting the molecular constitution of bodies, than the study of the fundamental notions and principles of that department of mechanics to which thermodynamics is especially related. (Gibbs 1902, p. ix)

The search for an appropriate *interpretation* of statistical mechanics which connects thermodynamics to statistical mechanics was a prime focus of both Gibbs and Boltzmann, even though their methods differed significantly.

This problem-situation – the search for molecular foundations for thermodynamics – remains relevant today. For instance, Callender (2001, p. 540) notes that “kinetic theory and statistical mechanics are in part attempts to explain the success of thermodynamics in terms of the basic mechanics.” More recently, Frigg and Werndl (2021) argues that this problem-situation is tied to a demand for

²⁶ See e.g. Seidenfeld (1986) and the references therein.

²⁷ The transition from Boltzmann to Gibbs is itself an interesting chapter in the history of statistical physics, but will be left, hopefully, to future work.

explanation, namely *why* and *how* statistical mechanics relates to thermodynamics at all. To ignore this problem-situation is akin to quietism about this relationship, and:

While practitioners may find it expedient to avoid the issue in this way, from a foundational point of view quietism is a deeply unsatisfactory position because it leaves the relation between [statistical mechanics] and [thermodynamics] (or indeed any macroscopic account of a system's behaviour) unexplained. (Frigg and Werndl 2021, p. 7)

This highlights a *need* to explain why thermodynamics is related to statistical mechanics, and that is precisely what the problem-situation is about. Indeed, Jaynes (1957, p. 620) situates his work in terms of this problem-situation as well. Suffice to say, this problem-situation is not an arbitrary one, but one that has motivated the foundations of statistical mechanics and plays a crucial role in its history and development.²⁸

As is well-known, the Gibbsian approach and its associated entropy face several conceptual troubles in attaining this goal. Recall the unclear nature of the fictitious ensembles and how to interpret the probabilities provided by the approach, as well as issues such as the requirement to assume various physical hypotheses. For instance, ergodicity or metric transitivity²⁹ is typically introduced as a founding assumption³⁰ in statistical mechanics textbooks to connect measured (i.e. time) averages to expectation values over phase space. However, there are some crucial

28 Of course, it may be the case that there are other problem-situations beyond the above. Wallace has recently argued that “it may be misleading to regard statistical mechanics itself as itself wholly or primarily a conceptual underpinning for thermodynamics (Wallace 2015, 286).” Instead, he opts for an approach that emphasizes an understanding statistical mechanics as used in contemporary practice, “concerned little with providing a foundational underpinning for the general principles of thermodynamics (Wallace 2015, 292)”, *contra* Frigg and Werndl above. This may well be true, and it is not the goal of this paper to argue for the “one true problem-situation” for statistical mechanics and entropy (if one should even exist). There is no reason why degeneration must be understood only in terms of some problem-situations and not others: we could always perform the analysis of degeneration in terms of Wallace’s proposed problem-situation. This raises an interesting further question of when and how problem-situations evolve and change. There could very well be an analysis of degeneration of problem-situations themselves, though I will leave that for future work. In any case, to my knowledge, Wallace’s proposal remains a minority view. The goal of searching for the molecular foundations of thermodynamics remains generally accepted as historically and conceptually important. Many thanks to an anonymous reviewer for pointing this out.

29 Roughly, ergodicity (or what Jaynes referred to as “metric transitivity”, and also known elsewhere as “metric indecomposability” (see Sklar 1993, 165) is the idea that the phase space for a system is such that a phase point’s trajectory is not confined only to one sub-region – another way to think about it is that the phase point’s trajectory traverses the entire volume of phase space.

30 See Frigg and Werndl (2021) for an overview of the status of this principle and related notions.

issues with this assumption: for instance, systems relevant to statistical mechanics are not always ergodic.³¹ Jaynes does discuss some of these worries, such as the appropriate interpretation of statistical mechanical probabilities and the requirement for Gibbsian statistical mechanics to include seemingly arbitrary “physical hypotheses” such as ergodicity or an *a priori* principle of indifference (Jaynes 1957, p. 621). However, the question of whether his paper contributes in a significant way to this problem-situation remains. Does his proposed interpretation of entropy-as-ignorance solve the issues faced by this problem-situation? Does it provide a clearer understanding of the issues above, or does it obfuscate the issues at hand?

The stated motivations for his paper (Jaynes 1957, p. 261) suggest preliminary grounds for concern about superfluity. Jaynes claims that his primary motivations are (i) bringing in new mathematical machinery to statistical mechanics, and (ii) the notion that information theory is “felt by many people to be of great significance for statistical mechanics”, although “the exact way in which it should be applied has remained obscure.” But these motivations do not help with respect to the problem-situation above. Jaynes also does not specify any concrete problem-situation relevant to statistical mechanics, only mentioning the above issues in passing.

4.2 Assessing the Depth of Jaynes’s Claims: Superfluity

The distinction between subjective and objective interpretations of probabilities was presumably an attempt by Jaynes to address the issue with interpreting the probabilities prescribed by Gibbsian statistical mechanics. This is, of course, a real interpretative issue with statistical mechanics. As mentioned, we can distinguish between interpreting probabilities as worldly objective chances, or subjective degrees of belief about events (which may or may not be subject to further rational constraints); the usual debate ensues as to which interpretation is appropriate.³² However, regardless of the result of *that* debate, Jaynes’s *actual* proposal with regards to MAXENT simply does not rely on a choice between them. As I see it, Jaynes’s proposed “subjectivist statistical mechanics” is simply generalization for generalization’s sake.

The discussion of the subjective/objective distinction is *superfluous* in the context of Jaynes’s paper. To see the irrelevance and superfluity of that distinction, consider that concepts like *information* and *uncertainty*, and what is sometimes confusingly called “knowledge” or “our knowledge” about something, in information theory, are in fact neutral between the two interpretations of probability (contrary to folk usage of these terms). The Shannon entropy is a formal quantity

³¹ See Earman and Rédei (1996).

³² For a more detailed discussion of the general positions one might take, see Sklar (1993, p. ch. 3).

that tracks the flatness of *any* probability distribution (be it a distribution for objective chances or degrees of belief): the more peaked it is, the more information (and less entropy) it contains. Indeed, looking at its use in communications, the relevant distributions involved are typically distributions of objective *frequencies* (e.g. of letters, words and so on), not *degrees of belief*. Unless one is understandably tricked by the occurrence of subjective-sounding words like “surprise”, “information” (in the sense that it informs *someone*), “uncertainty” and “knowledge”, the notion of information is, I claim, neutral between the objective and subjective interpretations of probability.

And why should it? Information theory is an extremely useful tool for our everyday communications, but it is ultimately a mathematical tool, without explicit metaphysical import. The distinction between objective and subjective interpretations of probability, on the contrary, is clearly a metaphysical one. As Jaynes notes: “The theories of subjective and objective probability are mathematically identical”, though they differ conceptually (Jaynes 1957, p. 622). So it is for the information-theoretic (Shannon) entropy. Even though common introductions gloss it as a measure of “uncertainty”, this does not force a subjective interpretation of probability onto the Shannon entropy.

If so, the MAXENT proposal – which simply requires maximizing the Shannon entropy of any probability distribution, over and above other physical constraints – is likewise *neutral* between interpretations. We can certainly *choose* to interpret MAXENT in a subjectivist way as Jaynes did. Since we are considering only probability distributions as degrees of belief, maximizing entropy is akin to adopting the “flattest” distribution of degrees of belief regarding a certain class of events given the available constraints. But we can also consider probability distributions as objective chances, in which case the MAXENT proposal becomes one in which we postulate that the probabilistic behavior of systems simply acts in a way that maximizes entropy given the constraints.

Jaynes seems to see the latter as unpalatable and the former acceptable: he mentions how his subjectivist proposal avoids “arbitrary assumptions” (Jaynes 1957, p. 630) or “physical hypotheses” (Jaynes 1957, p. 621) several times. But why should physical hypotheses be avoided or labelled arbitrary in the field we call *physics*? Ergodicity might have its own conceptual issues and concerns over applicability, but it is surely a valid hypothesis to be considered and debated, rather than dismissed seemingly *a priori* as one would in Jaynes’s approach. This is especially since ergodicity allows us to reproduce much of the physics we care about at the macroscopic level. And, at the very least, we are making a claim about the system’s actual behavior (which may obtain or otherwise) and why it fits the predictions we make about it in our theory. Compare this to the subjectivist proposal, in which the theory of statistical mechanics no longer describe the dynamics

of the chances of events occurring on phase space, but merely our degrees of belief about those events occurring. As Albert famously quipped,

Can anybody seriously think that our merely being *ignorant* of the exact microconditions of thermodynamic systems plays some part in *bringing it about*, in *making it the case*, that (say) *milk dissolves in coffee*? How could that *be*? What can all those guys have been *up to*? (Albert 2000, p. 64)

Is the subjectivist proposal really any less arbitrary when it comes to connecting our physical theories to the world? Jaynes does not elaborate. I do not want to adjudicate the debate here, though it suffices to say that the interpretation of probabilities is simply superfluous to the actual MAXENT proposal – the proposal itself, as a piece of mathematics, is independent of interpretation. Since both interpretations will inevitably reproduce the same mathematics (and hence the same equations), both interpretations inevitably rise and fall together.

There is little reason to think that Jaynes meant the MAXENT proposal to be much more than just a useful piece of mathematics that can help us compute and make predictions in a more tractable fashion, for it seems that his discussion of MAXENT entirely brackets off the issue of interpretation. If that is the case, however, the question of interpreting probabilities in statistical mechanics *does not even arise*. The actual goal of the paper is not about interpretation or the metaphysics of statistical mechanics. In turn, the question of interpreting the thermodynamic entropy, defined over these probabilities *about the system*, does not arise. Rather, MAXENT is a proposal concerning convenient prediction and computation. Jaynes writes:

Although the principle of maximum-entropy inference appears capable of handling most of the prediction problems of statistical mechanics, it is to be noted that prediction is only one of the functions of statistical mechanics. Equally important is the problem of interpretation; given certain observed behavior of a system, what conclusions can we draw as to the microscopic causes of that behavior? To treat this problem and others like it, a different theory, which we may call objective statistical mechanics, is needed. (Jaynes 1957, p. 627)

The MAXENT proposal is here just a convenient proposal for arriving at the computations required for prediction problems. Hence Jaynes claimed that adopting the “subjective point of view” and MAXENT for predictions serves a “great practical convenience”. But if we were to press Jaynes on the interpretation and metaphysics of statistical mechanics, we would still need “objective statistical mechanics”:

In the problem of interpretation, one will, of course, consider the probabilities of different states in the objective sense; i.e., the probability of state n is the fraction of the time that the system spends in state n . (Jaynes 1957, p. 627)

Jaynes’s take on the interpretation of probabilities *about the actual physical system* remains an objectivist one – and one seemingly adopting some version of the

ergodic hypothesis he claimed to have eschewed!³³ This goes to show that the probabilities prescribed by statistical mechanics about actual systems – and their interpretations – are not even in question here in Jaynes’s paper, since his proposal is supposed to be one concerning “subjective statistical mechanics”, rather than “objective statistical mechanics”. If we are only interested in prediction and computation, all we need is the *mathematical* MAXENT proposal, and a formal proof that it does in fact recover the equations we want. The ability for the MAXENT proposal to shorten and speed up the derivations of certain equations (as Jaynes shows in the paper) is, by and large, not in question here. Yet there is also no need to provide an interpretation for MAXENT and the notion of entropy involved in that case. There is no more need to interpret the mathematical shortcuts that one takes, any more than one needs to justify and interpret the algorithms behind WolframAlpha when one takes a shortcut with their integrals. All this renders Jaynes’s insistence on packaging the MAXENT proposal with a choice of interpretation for both probabilities and entropy confusing.

Furthermore, since we are not tackling the *actual* issue of how to interpret the probabilities assigned to the states of the actual system, the subjectivist MAXENT package is not even *relevant* to the original problem-situation. In other words, the insistence on providing an information-theoretic interpretation of entropy – replacing the previous thermodynamic and physical interpretation via the second law of thermodynamics and notions of reversibility/irreversibility of actual processes – is simply unjustified because the MAXENT proposal has no real need for interpretation.

No other genuine argument for the information-theoretic interpretation can be found in his paper. He starts off with a *proviso*:

The mere fact that the same mathematical expression $-\sum p_i \log p_i$ occurs both in statistical mechanics and in information theory does not in itself establish any connection between these fields. This can be done only by finding new viewpoints from which thermodynamic entropy and information-theory entropy appear as the same *concept*. In this paper we suggest a reinterpretation of statistical mechanics which accomplishes this, so that information theory can be applied to the problem of justification of statistical mechanics. (Jaynes 1957, p. 621)

As I have shown, information theory is not ultimately applied to the *justification of statistical mechanics*. That project requires interpreting statistical mechanics and

³³ As Frigg and Werndl (2021, p. 8) puts it, ergodicity holds “if for all measurable functions the infinite time average is equal to the ensemble average for almost all initial conditions”. Roughly, an ergodic principle states that the time-averages of some parameter is equal (in the infinite time limit) to the expectation value of that parameter (and hence the probability of that parameter having some quantity is equal to that amount of time the system spent in the region of phase space with that quantity, as Jaynes says here).

the metaphysics within it (e.g. about whether swarms of particles can actually recover the macroscopic description). Despite Jaynes's claim that he is proposing a reinterpretation of statistical mechanics, he does not succeed in doing so – that is all left in the “objective statistical mechanics” side of things, which he chose to downplay. Jaynes's proposal is the adoption of new mathematical tools for computing predictions in statistical mechanics, which does not force any interpretation at all. In any case, such interpretations have no real import for the actual issue of the problem-situation, that of interpreting the probabilities attached to events themselves. It is important to note that Jaynes does not specify an alternative problem-situation either. Instead, he simply asserts:

Since $[-\sum p_i \log p_i]$ is just the expression for entropy as found in statistical mechanics, it will be called the entropy of the probability distribution p_i ; henceforth we will consider the terms “entropy” and “uncertainty” as synonymous. (Jaynes 1957, p. 622)

But he has not yet shown that the thermodynamic entropy, i.e. “entropy”, and the information-theoretic entropy, i.e. “uncertainty”, are the same as a matter of interpretation, because the paper is not at all concerned with interpretation and “objective statistical mechanics”, only prediction and “subjective statistical mechanics”.

In sum, Jaynes's discussion of interpretative issues is *superfluous*. Jaynes has added unnecessary terminology from information theory and confused these new concepts with old questions without actually addressing any of the old questions from the original problem-situation. He has generalized for generalization's sake. The insistence on interpreting entropy as information-theoretic ignorance in a subjectivist sense, defined over distributions interpreted as degrees of belief, is likewise superfluous. As Denbigh correctly notes, “Jaynes' remark [on interpreting entropy in a subjectivist manner], though undoubtedly illuminating in a certain sense, is quite superfluous to the actual scientific discussion (Denbigh 1990, p. 111)”.

4.3 Assessing the Depth of Jaynes's Claims: Authoritarianism

Jaynes also displays authoritarianism when insisting on treating statistical mechanics as a general means of prediction, apparently viewed through subjectivist lens.

Authoritarians introduce new concepts into a line of inquiry without justification, while ignoring the problem-situation and heuristics which led us to those concepts. This is an important issue in Jaynes's paper, since the problem-situation at hand is barely specified. No details about the issues facing “objective statistical

mechanics” or the Gibbsian approach are presented. Instead, Jaynes presents the information-theoretic interpretation, the subjectivist interpretation, and the maximum entropy principle, as though they must be taken altogether.

Jaynes proclaims that in

freeing [statistical mechanics] from its apparent dependence on physical hypotheses of the above type, we make it possible to see statistical mechanics in a much more general light. (Jaynes 1957, p. 621)

Throughout the paper, Jaynes insists that the subjectivist approach is necessary for approaching the prediction issue. However, two questions arise. First, why the downplaying of “physical hypotheses” used by “objective statistical mechanics” and why do we need to “free” statistical mechanics from them? Second, why the focus on prediction and information theory, and the downplaying of the importance of interpretation? Both questions are unanswered.

To the first question, Jaynes demands that a satisfactory theory connecting microscopic to macroscopic phenomena should, among other things, “involve no additional arbitrary assumptions (Jaynes 1957, pp. 620–621)”. He notes a worry that this condition might be too severe since, rightfully, “we expect that a physical theory will involve certain unproved assumptions, whose consequences are deduced and compared with experiment (Jaynes 1957, 621).” However, his response to this worry is unsatisfactory. After listing some additional assumptions historically used in statistical mechanics to ensure empirical adequacy, he notes that

with the development of quantum mechanics the originally arbitrary assumptions are now seen as necessary consequences of the laws of physics. This suggests the possibility that we have now reached a state where statistical mechanics is no longer dependent on physical hypotheses, but may become merely an example of statistical inference. (Jaynes 1957, 621)

However, the fact that *some* arbitrary assumptions eventually come to be explained by quantum mechanics does not entail that statistical mechanics is (or should be) free of *all* physical hypotheses. The possibility that this *could* be possible is not a good argument for thinking that this *is* in fact the case (which would be necessary for him to argue so strongly against the use of physical hypotheses). Furthermore, even granting that this were true, the inference from this to statistical mechanics becoming “merely an example of statistical inference” is an unexplained leap as well.

In short, Jaynes believes that these physical hypotheses are undesirable – for instance, he spends some time claiming that metric transitivity is not needed if we adopt the MAXENT principle (Jaynes 1957, p. 624). However, he never provides an adequate reason for *why* we should not adopt *any* physical hypotheses about the systems we are studying.

He focuses instead on how MAXENT can help us do away with these hypotheses. But it is not clear that it does – MAXENT merely shifts our attention away from whether those hypotheses hold. Jaynes notes that adopting MAXENT is *in fact* akin to adopting ergodicity – except about our own degrees of belief about the system's behavior, rather than about the actual system's behavior:

Even if we had a clear proof that a system is not metrically transitive, we would still have no rational basis for excluding any region of phase space that is allowed by the information available to us. In its effect on our ultimate predictions, this fact [i.e. MAXENT] is equivalent to an ergodic hypothesis, quite independently of whether physical systems are in fact ergodic. (Jaynes 1957, p. 624)

Is the system *really* ergodic? And is ergodicity needed to derive the equations concerning those systems' behavior? Jaynes's proposal has two options: one is to say nothing at all – an unsatisfactory answer. Another option is to reply: the MAXENT proposal says that you should have degrees of belief matching the situation where the system is ergodic (as the quote above suggests). But that means I ought to believe that the system is ergodic after all, i.e. believing the *physical hypothesis of ergodicity*. Yet that was the original issue in our problem-situation: we want to know whether ergodicity is necessary for the statistical mechanical system to behave in accordance with our observations. Either MAXENT is irrelevant to our problem-situation, or it adds nothing new. Old questions remain.

The original problem-situation has been neglected. Yet, we are made to believe that these questions are to be ignored in favour of the new proposal – MAXENT, information theory, subjectivism – without justification for why that should be so. This is a case of authoritarianism.

Turning to the second question: as discussed above, the founding fathers of statistical mechanics were concerned first and foremost with the interpretative issues – how do we connect the particles or systems of statistical mechanics to the bulk macroscopic behavior we find in thermodynamics? Of course, that is not to say that prediction has no role to play in statistical mechanics. However, it is strange to ignore a core tenet of statistical mechanics, which seems like what Jaynes has done here. Reading the paper, one gets the impression that prediction holds supreme place in statistical mechanics. Interpretation seems to be an afterthought. But prediction goes hand in hand with interpretation – to predict the behavior of the system we must understand what the system *is*, and that is a matter of interpretation. As I have argued in Section 4.2, Jaynes's paper is completely divorced from interpretative issues. In this respect his paper is authoritarian: it ignores the problem-situation of statistical mechanics, such as the importance of interpretative issues.

The introduction of information theory, and the shift in focus on statistical mechanics as a general tool of statistical inference, is likewise authoritarian. Jaynes offers no reason for adopting information theory – we are told that the Gibbs entropy *can* be interpreted as the Shannon entropy, and that “the development of information theory has been felt by many people to be of great significance for statistical mechanics (Jaynes 1957, p. 621)”. Likewise, we are not told why statistical mechanics should be a general tool of statistical inference, freed from physics, where “the usual rules are thus justified independently of any physical argument, and in particular independently of experimental verification (Jaynes 1957, p. 620).” These are all core tenets of the MAXENT proposal, but they remain unjustified.

In conclusion, Jaynes’s paper falls afoul of both superfluity and authoritarianism. With respect to methodological depth, then, it was a degenerative piece of work. Since the key transition of entropy from a concept concerned with thermodynamics and actual physical systems to a concept concerned with ignorance and our knowledge of said systems occurred here, this shift is a degenerative one as well.

Jaynes’ paper changed the trajectory of the entropy concept. For instance, by appearing as though the paper presented an interpretative package, despite the actual proposal not needing one, the paper introduced confusion to the actual interpretative issues. The interpretative package of information-theoretic entropy and subjectivism became adopted as an answer for the interpretative issues associated with “objective statistical mechanics” instead. Recall the Bekenstein quote: a mere 20 years later the information-theoretic interpretation has escaped from “subjective statistical mechanics” into “objective statistical mechanics”, with *thermodynamic* entropy (defined over probability distributions of the microstates of the *actual* systems) being interpreted as ignorance or uncertainty. Three years later, Hawking would simply assert: “an intimate connection between holes (black or white) and thermodynamics [...] arises because information is lost down the hole (Stephen 1976, p. 197).” Degeneration has occurred.

4.4 Content-Oriented Degeneration

It is worthwhile to conclude by briefly considering the content-degeneration of the MAXENT proposal under the generalized Lakatosian account I have developed in Section 3. It seems to me that MAXENT is both theoretically and empirically degenerative on this account.

Recall the definitions: being theoretically progressive refers to a succeeding theory predicting more novel facts compared to its predecessor. Being empirically progressive refers to the excess empirical content of this succeeding theory

actually leading to the discovery of new facts, thereby corroborating the new theory's novel predictions.

As Jaynes notes, nothing new is added in terms of theoretical progress, because “subjective statistical mechanics” will recover *exactly the same predictions* as “objective statistical mechanics”:

Conventional arguments, which exploit all that is known about the laws of physics, in particular the constants of the motion, lead to exactly the same predictions that one obtains directly from maximizing the entropy. (Jaynes 1957, p. 624)

And that

the subjective theory leads to exactly the same predictions that one has attempted to justify in the objective sense. (Jaynes 1957, p. 625)

This shows that no new predictions are provided by this proposal. The “new” proposals attached in Jaynes's papers are typically just new ways of doing the same calculations. For instance: Jaynes's treatment of Siegert's “pressure ensemble” is also merely a reworking of Siegert's own derivations, published just a year prior (Lewis and Siegert 1956). In short, the MAXENT proposal is theoretically degenerative. Furthermore, since there are no new predictions, there are no new predictions to corroborate. The proposal is empirically degenerative. This, of course, also adds to that sense of superfluity one gets when analyzing Jaynes's proposal in detail.

Overall, then, Jaynes's paper is degenerative *tout court* in terms of its place in statistical mechanics.³⁴ Given that it had such a huge influence on the current understanding of entropy as ignorance and uncertainty, especially in the field of contemporary black hole thermodynamics (Bekenstein and Hawking are typically known as the founding fathers of black hole thermodynamics), this current understanding must be re-assessed. Some have already begun this work. For instance, Wüthrich have recently argued that

the original argument by Bekenstein with its detour through information theory does not succeed in establishing the physical salience of the otherwise merely formal analogy between thermodynamic entropy and the black hole area, and so cannot offer the basis for accepting black hole Thermodynamics as “the only really solid piece of information”. (Wüthrich et al. 2018, pp. 219–220)

34 There is certainly room to argue for this paper's contributions in the context of other fields, given Jayne's and MAXENT's ubiquity outside of statistical mechanics. However, this is not Jayne's original goal.

Importantly, Wüthrich diagnoses the problem with Bekenstein's arguments as the failure to recognize that "Fundamental physics is about the objective structure of our world, not about our beliefs or our information", and that "information, one might argue, is an inadmissible concept in fundamental physics (Wüthrich et al. 2018, p. 217)." Given my analysis here, we can see why that is the case. The introduction of information theory by Jaynes to statistical mechanics was already superfluous to begin with. Others like Prunkl and Timpson (1903), recognizing the flaws with information-theoretic arguments for black hole thermodynamics, are already attempting to provide a defense of black hole thermodynamics *sans* information theory. A possibility of doing so further suggests – in agreement with my diagnosis here – that information-theoretic concepts may just have been superfluous to the discussion, and, to quote Wuthrich again, a "detour".

5 Conclusion

I have provided and motivated an extension to Lakatos's account of growth and degeneration from MSRP by appealing to P&R. This extension, in terms of superfluity and authoritarianism, enables a new dimension through which we may evaluate a piece of mathematical or scientific work, independent of the analysis in terms of theoretical and empirical progress or degeneration found in MSRP.

As proof of concept, I have evaluated Jaynes's proposal, a key transition point in the historical trajectory of the concept of entropy. I hope to have shown that my account does provide a novel means of assessing the degeneration or progress of this transition, by critically analyzing the aspects of his paper which exhibited superfluity and authoritarianism.

Some might object that my criticisms of Jaynes' proposal could have been made independent of the account of degeneration I have sketched here.³⁵ I agree that one *could* have arrived at these criticisms independent of my account, for there are likely many ways to arrive at the same conclusion I reached. However, that does not discount the fact that my account of degeneration *does* arrive at these criticisms, guided by the twin heuristics of superfluity and authoritarianism. I hope to have shown in this paper that this account provides us with a grip on the nature of these criticisms (as methodological ones) and motivates them in a conceptually clear fashion. This should give us a good reason to consider and adopt this account of degeneration regardless of whether there might be other ways to arrive at these criticisms. In any case, we should rejoice – not despair – when there are multiple

³⁵ Many thanks to an anonymous reviewer who raised this point.

ways of evaluating a problem, for this means we have more tools in our conceptual toolbox for analysis.

In my view, developing more tools for understanding how scientific and mathematical concepts degenerate or grow has natural affinities with an increasingly popular understanding of philosophy as conceptual engineering.³⁶ As Chalmers (2020, p. 4) writes, conceptual engineering is “the project of designing, evaluating, and implementing concepts”, where we consider not only what a concept *is*, but also what it *should be*. Developing new tools for identifying points of degeneration in a concept’s historical trajectory helps us evaluate a concept and consider alternative ways of designing and developing said concept.

This paper thus leaves behind a variety of fruitful directions, ripe for the picking by the hopeful conceptual engineer. For those who, like me, are puzzled by the concept of entropy: if we want to re-engineer and design a newer, better, conceptually clearer notion of entropy, we would do well to engage with – and dispel – other similarly degenerative transition points. For other philosophers, too, I believe the tools developed here can be used to assess concepts elsewhere: in science, mathematics, perhaps even philosophy itself. There remains much to be done.

References

- Albert, D. Z. 2000. *Time and Chance*. Cambridge: Harvard University Press.
- Bekenstein, J. 1973. “Black Holes and Entropy.” *Physical Review D* 7 (8): 2333–46.
- Bub, J. 2005. “Quantum Mechanics is About Quantum Information.” *Foundations of Physics* 35 (4): 541–60.
- Callender, C. 1999. “Reducing Thermodynamics to Statistical Mechanics: The Case of Entropy.” *The Journal of Philosophy* 96 (7): 348–73.
- Callender, C. 2001. “Taking Thermodynamics Too Seriously.” *Studies in the History and Philosophy of Modern Physics* 32: 539–53.
- Chalmers, D. 2020. “What is Conceptual Engineering and What Should it Be?” *Inquiry*, <https://doi.org/10.1080/0020174X.2020.1817141>.
- Chua, E. Forthcoming. “Does Von Neumann Entropy Correspond to Thermodynamic Entropy?” *Philosophy of Science* 88 (1): 145–68.
- Corfield, D. 1997. “Assaying Lakatos’s Philosophy of Mathematics.” *Studies in History and Philosophy of Science* 28: 99–121.
- Denbigh, K. 1990. “How Subjective is Entropy?” In *Maxwell’s Demon: Entropy, Information, Computing*, edited by H. Leff, and A. Rex. New Jersey: Princeton.
- Dougherty, J., and C. Callender. “Black Hole Thermodynamics: More Than an Analogy?” Preprint. <http://philsci-archive.pitt.edu/id/eprint/13195> (accessed March 01, 2021).

³⁶ See e.g. Haslanger (2000) or Chalmers (2020).

- Earman, J., and M. Rédei. 1996. "Why Ergodic Theory Does Not Explain the Success of Equilibrium Statistical Mechanics." *The British Journal for the Philosophy of Science* 47: 63–78.
- Friston, K., and K. E. Stephan. 2007. "Free-Energy and the Brain." *Synthese* 159 (3): 417–58.
- Frigg, R., and C. Werndl. 2021. "Can Somebody Please Say What Gibbsian Statistical Mechanics Says?" *The British Journal for the Philosophy of Science*, 72 (1): 105–29.
- Gibbs, J. W. 1902. *Elementary Principles of Statistical Mechanics: Developed with Special Reference to the Rational Foundation of Thermodynamics*. New Haven: Yale University Press.
- Goldstein, S., J. Lebowitz, R. Tumulka, and N. Zanghi. 2020. "Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics." In *Statistical Mechanics and Scientific Explanation*, edited by V. Allori, 519–81. Singapore: World Scientific.
- Kampis, G., L. Kvasz, and M. Stöltzner. 2002. *Appraising Lakatos: Mathematics, Methodology, and the Man*. Dordrecht: Kluwer.
- Hallett, M. 1979a. "Towards a Theory of Mathematical Research Programmes I." *The British Journal for the Philosophy of Science* 30 (1): 1–25.
- Hallett, M. 1979b. "Towards a Theory of Mathematical Research Programmes II." *The British Journal for the Philosophy of Science* 30 (2): 135–59.
- Haslanger, S. 2000. "Gender and Race: (What) Are They? (What) Do We Want Them to Be?" *Noûs* 34 (1): 31–55.
- Jaynes, E. T. 1957. "Information Theory and Statistical Mechanics." *Physical Review* 106 (4): 620–30.
- James, S. 1904. *Entropy, or, Thermodynamics from an Engineer's Standpoint and the Reversibility of Thermodynamics*. Constable: Westminster.
- Kiss, O. 2006. "Heuristic, Methodology or Logic of Discovery? Lakatos on Patterns of Thinking." *Perspectives on Science* 14 (3): 302–17.
- Ludwig, B. 1896/1995. *Lectures on Gas Theory*. New York: Dover.
- Lakatos, I. 1976/2015. In *Proofs and Refutations: The Logic of Mathematical Discovery*, edited by J. Worrall, and E. Zahar. Cambridge: Cambridge University Press.
- Lakatos, I. 1978. In *The Methodology of Scientific Research Programmes. Philosophical Papers Volume 1.*, edited by J. Worrall, and G. Currie. New York: Cambridge University Press.
- Leng, M. 2002. "Phenomenology and Mathematical Practice." *Philosophia Mathematica* 10: 3–25.
- Lewis, M. B., and A. J. F. Siegert. 1956. "Extension of the Condensation Theory of Yang and Lee to the Pressure Ensemble." *Physical Review* 101 (4): 1227–33.
- Sklar, L. 1993. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge: Cambridge University Press.
- Marcel Brillouin, L. 1956. *Science and Information Theory*. New York: Academic Press.
- Mclrvine, E. C., and M. Tribus. 1971. *Energy and Information*. Also available at <https://www.scientificamerican.com/article/energy-and-information/>.
- Natsuume, M. 2015. *AdS/CFT Duality User Guide In Lecture Notes In Physics 903*. Tokyo: Springer.
- Prunkl, C., and C. Timpson. 1903. *Black Hole Entropy Is Thermodynamic Entropy*. Preprint. ArXiv: 1903.06276.
- Roach, Ty. N. F. 2020. "Use and Abuse of Entropy in Biology: A Case for Caliber." *Entropy* 22 (12): 1335.
- Robertson, K. 2020. "Asymmetry, Abstraction, and Autonomy: Justifying Coarse-Graining in Statistical Mechanics." *The British Journal for the Philosophy of Science* 71 (2): 547–79.
- Stephen, H. 1976. "Black Holes and Thermodynamics." *Physical Review D* 13 (2): 191–7.
- Seidenfeld, T. 1986. "Entropy and Uncertainty." *Philosophy of Science* 53 (4): 467–91.

- Stöltzner, M. 2002. "What Lakatos Could Teach the Mathematical Physicist." In *Appraising Lakatos: Mathematics, Methodology, and the Man*, edited by K. George, L. Kvasz, and M. Stöltzner. Dordrecht: Kluwer.
- Uffink, J. 2001. "Bluff Your Way in the Second Law of Thermodynamics." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32 (3): 305–94.
- Wallace, D. "The Necessity of Gibbsian Statistical Mechanics." Preprint. <http://philsci-archive.pitt.edu/id/eprint/15290> (accessed March 01, 2021).
- Wallace, D. 2015. "The Quantitative Content of Statistical Mechanics." *Studies in History and Philosophy of Modern Physics* 52: 285–93.
- Werndl, C. 2009. "Justifying Definitions in Mathematics: Going beyond Lakatos." *Philosophia Mathematica* 17 (3): 313–40.
- Wüthrich, C. 2018. "Are Black Holes about Information?" In *Why Trust A Theory?: Epistemology of Fundamental Physics*, edited by R. Dardashti, R. Dawid and K. Thebault, 202–23. New York: Cambridge University Press.