Eulerian Routing in Practice

Davide Rizza[1]

Abstract

The Königsberg bridge problem has played a central role in recent philosophical discussions of mathematical explanation. In this paper I look at it from a novel perspective, which is independent of explanatory concerns. Instead of restricting attention to the solved Königsberg bridge problem, I consider Euler's construction of a solution method for the problem and discuss two later integrations of Euler's approach into a more structured methodology, arisen in operations research and genetics respectively. By examining Euler's work and its later developments, I achieve two main goals. First, I offer an analysis of the role played by mathematics as a problem-solving instrument within scientific enquiry. Second, I shed light on the broader significance of well known contributions to the debate on mathematical explanation. I suggest that these contributions, which are tied to a localised explanatory context, achieve a greater relevance and attain a sharper formulation when they are referred to scientific enquiry at large, as opposed to its possible explanatory outcomes alone.

1 From Explanation to Problem-Solving

In the eighteenth century, the river Pregel divided the city of Königsberg into four land masses joined by seven bridges. The Königsberg bridge problem, which attracted Euler's attention, is the problem of determining whether the bridges could be visited along a trail.[1] Euler's treatment of this problem has attracted much attention among philosophers of science and mathematics, at least since Pincock (2007) discussed it as an instance of abstract mathematical explanation.[2]

---

[1] A trail is a route crossing each bridge exactly once. A trail may go back to its starting point, in which case it is called a circuit.

[2] An interesting discussion of the problem connected with mathematical practice appears in Gaifman (2005, pp. 365–367).

✉ Davide Rizza
  d.rizza@uea.ac.uk

[1] School of Politics, Philosophy, Language and Communication Studies, Faculty of Humanities, University of East Anglia, Norwich NR4 7TJ, UK

Philosophers concerned with explanation have included the Königsberg bridge problem in their range of paradigmatic examples [see e.g. Lange (2013), Lyon (2012)], they have examined it in its historical context (Räz, 2018) or they have disputed its explanatory role (Jansson & Saatsi, 2019; Kuorikoski, 2021, p. 194)]. In this paper, I examine the Königsberg bridge problem from a standpoint that is independent of explanatory concerns (Sects. 2, 3, 4, 5). Then, I use the newly articulated standpoint to highlight the broader significance of certain contributions to the debate on mathematical explanation (Sect. 6). Before actually carrying out these tasks, I briefly clarify what standpoint I am going to take and how its full articulation can illuminate the significance of contributions focussing on mathematical explanation.

I take the standpoint recommended in the following remark made by Hasok Chang:

> A serious study of scientific practice must be concerned with what it is that we actually do in scientific work. This requires a change of focus from propositions to activities (Chang, 2011, p. 208).

Discussions of mathematical explanation typically exhibit a propositional tendency, which may be motivated by their special aims. Such discussions take a solution to a particular problem and the mathematical results leading to that solution as given: their key interest pertains to the distinctive manner in which the given elements are connected. Their focus is on settled facts and results, not on the activities directed at the construction of methods that can be adopted to solve open problems. Although a narrower explanatory focus is natural and legitimate for a highly circumscribed purposes (e.g. spelling out what a mathematical explanation is), it involves a deletion of the unfolding of enquiry that precedes the possible availability of explanations and upon which such availability depends.

It seems to me important to pay proper attention to the unfolding of scientific enquiry and the activities it consists of, which are in essence problem-solving activities, for two reasons. On the one hand, looking at problem-solving activities in progress brings to light many subtle aspects of the application of mathematics that simply cannot emerge if only the circumstances of explanation are attended to. On the other hand, a clear understanding of the role played by mathematics in scientific problem-solving enables a better appreciation of certain insights coming from discussions of mathematical explanation. These discussions identify important aspects of the application of mathematics, whose scope far exceeds the single moment of explanation. The broader significance of contributions to mathematical explanation cannot however be appreciated if an account of the problem-solving activities preceding the explanatory moment is missing. In fact, the absence of such an account forces insights into the structure of scientific enquiry into metaphysical formulations that partly conceal their significance.

I shall return to the last remarks in Sect. 6, where I spell them out in connection with three especially valuable contributions to the discussion of mathematical explanation, namely Pincock (2007, 2015a) and Baker (2017). As I have remarked, the analysis of Sect. 6 becomes possible only after a structured study of mathematical-problem solving is in place. Starting from Sect. 2, I intend to

provide such a study in connection with the Königsberg bridge problem. The latter problem provides a useful, but by no means self-contained, starting point. Euler's work on the Königsberg bridge problem was meant to supply a method capable of solving a family of what we would today call routing problems. Because of this, Euler's work points, as it were, beyond itself. In an expected manner, it points to its later integration into larger and more difficult routing problems, one of which is examined in Sect. 3. In a less expected manner, Euler's routing method has proved helpful in the field of genetic sequencing, as will be seen in Sect. 4.

It is important to emphasise that none of the applications discussed in Sects. 3 and 4 amounts (with the exception of especially fortunate cases) to a direct use of Euler's method yielding the desired results. Rather, this method is systematically coordinated with other mathematical techniques to confront a variety of more or less unwieldy circumstances. This is an important point because it suggests that the successful application of mathematics does not always amount to a straightforward matching between mathematical structure and empirical configuration but can in fact result from a mutual adaptation of plastic symbolic resources and controllable empirical features.[3]

This process of adaptation, which at times requires a reconstruction of mathematical method and at times requires novel empirical interventions (a striking instance will be discussed in Sect. 5) is the subject of the next four sections. One useful way of summarising their contents is to relate them to Chang's definition of an epistemic activity as:

> a coherent set of mental or physical actions (or operations) that are intended to contribute to the production or improvement of knowledge in a particular way, in accordance with some discernible rules (Chang, 2011, p. 209).

In terms of the definition above, Sects. 2, 3, and 4 focus on epistemic activities that are mental operations capable of improving knowledge by solving problems that bring practical goals within reach or make desired information accessible, in accordance with rules provided by mathematical methods. Thus, 'scientific problem-solving by mathematical means' may be, for the purpose of this paper, taken as synonym with 'epistemic activity'. The specific epistemic activities examined in the following three sections are: (i) the construction of a problem-solving method, which I shall discuss with reference to Euler (1741) (in Sect. 2);[4] (ii) the integration of a method into a broader problem-solving strategy, which is required to make Euler's result applicable to an actual routing problem, namely the optimal scheduling of street sweeping (investigated in Sect. 3); (iii) the adaptation of a problem to a method and the corresponding modulation of the method to suit the given problem (discussed

---

[3] The latter fact cannot stand out in explanatory accounts of the Königsberg bridge problem, where a simple setting directly supports a sharp conclusion.

[4] Räz (2018) also offers an insightful study of Euler's original paper. I suggest in the next section that Räz's analysis is hindered by the decision to read Euler's work along explanatory lines foreign to it. The epistemic activities conspicuous in Euler's paper are then lost sight of.

in Sect. 4, in connection the use of Euler's result in genetic sequencing). In Sect. 5, I briefly look at the impact of technological advances on the applicability of mathematical problem-solving strategies, focussing on DNA computing.

The paper is concluded by Sect. 6, which has already been summarised, and a final overview in Sect. 7.

2 Traversability

The Königsberg bridge problem was originally solved in Euler (1741). More precisely, Euler took that problem to suggest a general one: to give necessary and sufficient conditions under which a configuration of localities and pathways joining them is traversable along a trail (as defined in Fn. 1). I shall refer to this problem simply as the traversability problem for a connected configuration.[5] This is the actual problem solved in Euler (1741). Once one knows how to solve the general traversability problem, it is easy to apply Euler's solution method to the special Königsberg bridge configuration and find that it admits no trails.

Euler (1741) shows, in a particularly transparent way, three important phases involved in the construction of a problem-solving method by mathematical means. These phases are characterised by the increasing substitution of inference for direct checks based on the problem's data. In plainer terms, mathematical argument is introduced to avoid extensive processing of data. In the initial stage, Euler considers a situation where everything has to be checked and nothing is inferred: here one deals with data all the time, without establishing any inferential connection between them and the existence of a solution. In the intermediate stage, a targeted selection of the problem's data makes it possible to establish the desired connection. In the final stage, a theoretical refinement of the intermediate stage streamlines the intermediate, inferential refinement.

Euler's study of the Königsberg bridge problem can thus be viewed as a simple yet instructive illustration of the central role played by mathematical considerations in guiding pertinent selections and uses of a problem's data. Here pertinence must be referred to a preliminary goal (solving the problem). It is only after the epistemic activity of method construction is carried out that some data stands out as relevant. Its relevance is not an absolute feature, but it is relative to a method and a goal.

The picture of method construction just sketched can now be brought to life by looking at Euler (1741). Because a good summary and a thorough analysis of this paper are already offered in Räz (2018), I would like to clarify what the motivation for turning to Euler's text again is. Räz's aim is to discuss Euler's paper from an explanatory point of view (Räz, 2018, p. 335). For this reason, his analysis follows a trajectory that is very different, although not altogether unrelated, to that pursued here. In particular, Räz helpfully isolates three stages in Euler's paper, but he relates them to three distinct attempts at explaining the solved Königsberg bridge problem.

---

[5] A configuration that is not connected includes at least two localities that are not linked by any path. In this case the configuration cannot be traversable.

In Räz, these stages illustrate antagonistic explanatory strategies. A comparison of their explanatory merits motivates Räz to suggest a characterisation of explanatory power in terms of the selection of relevant information and the elimination of redundant information.

I think Räz correctly isolates three stages in Euler's paper, and I single out the very same stages in the discussion to follow. However, I provide a substantially different reading of these stages, as three interconnected phases of method construction, as opposed to antagonistic explanatory strategies. The reading I propose seems to me in line with the goals of Euler (1741), which are not explanatory. I also abandon the notion of relevant information, which Räz adopts, and replace it with relevance relative to a method and a goal. The notion of relevance so relativised is clear: data or information is relevant if a method explicitly selects it to solve a problem. This notion of relevance does not require further elaboration[6] and is only provided as a useful reminder that 'relevance' is, from the point of view of practice, the label assigned to data that can be subjected to specific use. Data does not display its relevance upon inspection, but only because it can be included into successful epistemic activities.

With these clarifications in the background, I now turn to Euler (1741) and isolate three stages within Euler's paper, which function as steps within a continuous process of method construction. In the initial step, Euler observes that one way to solve the traversability problem is to carry out an exhaustive search for trails. This can be concretely done on a journey or, symbolically, on a map. The strategy works and has the advantage of generating all trails, if any exist.

Euler, however, notes that exhaustive search is *de facto* inapplicable to very large configurations. It may also produce a large amount of information that is irrelevant to the goal at hand: if a trail exists, one does not need to acknowledge it only after describing every non-trail. Euler observes that a simpler approach is to look for a criterion that determines traversability.[7] The criterion is chosen, at least in the first instance, in such a way that it does not have to exhibit a trail when one exists. The latter goal is desirable but it may be convenient to renounce it if a less demanding goal may be attained in a more expeditious way. It is always possible to turn to the problem of constructing a trail later, as Euler himself does.

Euler's approach implicitly highlights one further issue with exhaustive search: beside possibly generating much irrelevant information, exhaustive search is an undiscerning strategy. It does not respond to empirical advantages. If, for instance, a configuration had one locality x joined to only one other locality by a single pathway,[8] one might only check the trails starting at x (clearly x will not be an intermediate locality along any trail). A complete search is not responsive to this sort of

---

[6] Räz has to try and unpack his notion of relevance in terms of logical strength or computational complexity [see Räz (2018, pp. 338–342)].

[7] '[...] in aliam inquisivi, quae plus non largiatur, quam ostendat, utrum talis cursus institui queat an secut' (Euler, 1741, pp. 129–130)

[8] Here and in the rest of the section, I use 'pathway' informally, as a counterpart to the formal term 'edge' from graph-theory. I am therefore not talking about graph-theoretic paths.

opportunity. It is also difficult to make it responsive, unless one has been able to select data that help curtail the search. This is precisely what Euler sets out to do in the second stage of problem-construction.

His goal is to move toward an existence criterion that will select some data and make use of it to anticipate the consequences of traversing a configuration. One way of identifying the data that it might be useful to select is to refer it to the set goal, i.e. traversability. Euler thus supposes that a solution to the traversability problem is given and proceeds to look for any features it must possess. This is difficult to do without a suitable notation. Euler assigns letters to the localities in a configuration, thus obtaining a uniform representation for the terms of the traversability problem. Localities are one-letter words, pathways joining them are two-letter words and a solution is a $(n + 1)$-letter word, if there are n pathways.

A solution to the traversability problem is completely specified when a $(n + 1)$-letter word is given and its constituent letters occur the right number of times in the right order. Since an existence criterion is less demanding than the exhibition of a solution, it is plausible to try and drop some of the information that uniquely determines a solution. The length of a solution is easily available, so there is no reason to drop it. If the complete, ordered sequence of letters was known, then the number of occurrences of each letter would be known. One could therefore drop order. Euler notes that the number of occurrences of the same letter in a solution can be computed.[9] Let this number be m_i for locality i. If there are n pathways and k localities, a trail exists if, and only if, $m_1 + \ldots + m_k = n + 1$.

The computations required to verify the last equality include a direct count of the number q_i of pathways emanating from each locality. This is all the data one has to select. Once the q_i are known, it can be decided whether a trail exist by further computation. The data supplied by the q_i associated with a configuration is all that is relevant, given Euler's method, to the set goal.

From a logical point of view, this characterisation of traversability cannot be improved upon. Any other characterisation will be equivalent to it. Nonetheless, the given characterisation only depends on whether an equality is true or not in a finite configuration, not on what values its sides take. Thus, it may be possible to truncate the criterion just obtained: Euler does precisely that:[10] he trades off arithmetical computations required in each specific traversability problem for combinatorial

---

[9] The computation only requires knowing the distinctive feature of a solution, namely the fact that pathways already traversed cannot be used again. Thus, for instance, if $2k$ pathways emanate from a locality A and a solution exists, then A occurs k times in it, if it is not the starting point of the trail, and it occurs $k + 1$ times otherwise. This is because, if A is always an intermediate stage in the trail, then any one of its occurrences uses up two pathways and, since a solution must visit every pathway only once, it must return to A exactly k times. If A occurs at the beginning of the trail, then one pathway is used to leave it. Every subsequent, intermediate transition uses up two pathways: $k - 1$ such transitions will subsequently be made. Then A must be returned to using the last available bridge. Thus A occurs once at the beginning, once at the end, and $k - 1$ times in the middle of the trail: overall, it occurs $k + 1$ times.

[10] Räz (2018, p. 338) seems to suggest that Euler's truncated criterion is logically weaker than his preceding criterion. I don't think this is the case: one criterion only requires fewer computations, but both characterise traversability. Moreover, if the additional computations are unnecessary when traversability fails, they provide information that serves the purpose of constructing a solution when one exists.

considerations that, once made, license a general argument. Euler is thus able to conclude that a configuration is traversable if, and only if, q_i is even for $i = 1, \ldots, k$ or exactly two of the q_i are odd. It is worth stressing that this characterisation of traversability does not reduce the data previously required but only the number of computations involving them.

What now emerges is a clear outline of the process that leads to the construction of a traversability criterion in Euler (1741). The process begins with a situation in which no data can be selected. The traversability goal is then employed as a guide to data selection. A traversability criterion emerges. A closer analysis of the way relevant data is used by the criterion leads to a computationally less expensive variant.

Once constructed, a mathematical method like Euler's traversability criterion is not typically employed to supply explanations and settle arbitrary traversability puzzles. Rather, it is involved in more articulate epistemic activities (Euler himself saw it as a contribution to non-metrical geometry). For this to happen in applications, the criterion cannot be wielded as an isolated result. In the first instance, it must be equipped with an algorithm that produces trails when they exist.[11] From now on, I shall refer to the complex of Euler's criterion and the associated algorithm by the expression 'Eulerian routing'.

Eulerian routing provides a simple mathematical methodology available to tackle many familiar problems: the one I shall turn to in the next section is the optimal scheduling of street sweeping. The main reason why it is worth looking at it is that, despite being similar to the Königsberg bridge problem, street sweeping presents empirical adversities that make Eulerian routing an ineffective methodology when taken on its own. The response to this situation is the integration of Eulerian routing into a broader methodology.

3 Routing

Eulerian routing may prove insufficient in the face of the adversities that occur when actual routing tasks are taken into account. In practice, if a trail can be determined, it may be necessary to travel through it within a prescribed time. If no trails exist, it will not be enough to record the fact and resign oneself to it, as would have been the case with the Königsberg bridge problem. Certain operations have to be carried out whether or not they can be routed on a trail. Finally, if there are distinct trails, the selection of one among them may not be a matter of indifference: some trails will produce costs that others save.

All of these issues occur in the optimal scheduling of street sweeping. On account of parking regulations, mechanical brooms can only sweep within prescribed times. Moreover, because what is swept by a single vehicle is usually a single city district,

---

[11] Euler (1741, p. 140) offers some remarks supporting this conclusion, but the first published algorithm that constructs a trail if one exists was devised by Carl Hierholzer in 1873. Hierholzer's algorithm is discussed in Sect. 4. See Fleischner (1990, pp. 30–32) for the German original as well as an English translation of Hierholzer's paper.

its street network may not be traversable in Euler's sense. Finally, specific trails, e.g. ones avoiding inconvenient turns (where the street side to be swept switches), are preferable because they allow a simpler operation of the broom.

These issues, which I shall collectively refer to as the empirical adversities of street sweeping, do not evidently suggest that Eulerian routing should be given up. They call for its adaptation to new circumstances and, in particular, for its articulation with additional methods and techniques that can be turned to when the straightforward, basic routing methodology fails. These observations can be made more concrete and more compelling by looking at the integration of Eulerian routing into the construction of optimal street sweeping schedules devised for the New York City Department of Sanitation by Bodin and Kursh.[12]

What is, in the present context, striking about the construction of an optimal schedule is that it can be naturally seen as the progressive unfolding of a problem-solving strategy whose successive steps are determined by typical empirical adversities. In the simplest case, the methodology coincides with Eulerian routing. The need for responses to adversities articulates Eulerian routing with additional algorithms and results.

To see how, it is necessary to clarify what the terms of the street sweeping problems are. The problem consists in assigning mechanical brooms routes that allow them to visit, within a specified time period, every car-free curb in a city district whilst avoiding inconvenient turns and, if necessary, minimising deadheading time (i.e. the time a mechanical broom travels without sweeping). It may also be desirable, and will be assumed in the following discussion, that a broom should travel along an optimal circuit (see Fn. 1 for a definition), so that it goes back to its depot at the end of the journey.

The routing component of the problem requires focussing on a street network: for the sake of terminological convenience, I shall refer to crossroads as nodes in the street network and to streets as directed edges joining nodes. Direction matters because some streets are one-way (a two-way street is formally conceived of as a pair of directed edges pointing in opposite directions).

If a city district is traversable along a circuit, the circuit is an optimal route for a mechanical broom. Only a small departure from Euler's original problem occurs, because the direction of travel now matters. Euler's traversability criterion must be revised to take it into account. Whenever an intermediate node is traversed, this is done through an incoming edge and a distinct outgoing edge. In a circuit, every node is intermediate: thus, incoming and outgoing edges must be balanced at each node. In this case, Eulerian routing along a circuit is possible.[13]

---

[12] The formal methodology is presented in Tucker and Bodin (1976) and an early computer-assisted implementation is described in Bodin and Kursh (1978).

[13] The circuit can then be determined using the routing criterion. The key idea is that leaving a starting node costs one (outgoing) directed edge, whilst traversing any intermediate node costs two directed edges (incoming and outgoing). A path leaving a directed edge will return to it, if a circuit exists. If such a path is not a complete circuit, other circuits can be glued to it for the reason just stated. The result is a complete circuit.

The important practical issue is to figure out what to do when a city district does not allow Eulerian routing. If $\mathcal{G}$ is the street-network of interest, the natural strategy is to force a circuit by minimal detours. More formally, one wants to add a minimal-length set of directed edges E to the original network $\mathcal{G}$ in such a way that the resulting, enlarged street-network $\mathcal{H}$ allows Eulerian routing.

The minimality constraint on E depends on the fact that no edge of the larger street-network $\mathcal{H}$ that is not already in $\mathcal{G}$ has to be swept. The adjunction of E to $\mathcal{G}$ now calls for mathematical resources that go beyond Eulerian routing. It is worth stressing that these resources play the fundamental role of telling the problem-solver what to do: they guide a particular activity.

In the first instance, a theorem about networks provides information on the way to choose the needed edges. The theorem states that, if an optimal collection of edges E exists, then it can be partitioned into paths from nodes of negative degree in $\mathcal{G}$ to nodes of positive degree in $\mathcal{G}$.[14] Only shortest paths linking such nodes are to be taken into account: Dijkstra's algorithm singles them out.[15] Finally, these paths are allocated through the solution of a transportation problem.[16] Substantially more than just Eulerian routing is needed.

The last remarks are of philosophical importance because they draw attention to the fact that the applicability of mathematics does not only depend on a suitable structural display on the part of the phenomena, but also on the plasticity of problem-solving strategies. This is to say that the applicability of a specific approach like Eulerian routing does not depend entirely on this strategy taken in isolation, but also, and crucially, on its possible articulations with other techniques and results. Thus, an application of Eulerian routing is not typically successful because it reduces to a 'one-shot' use of a single criterion. It is, in general, successful insofar as it can be embedded into a a progressively richer mathematical methodology capable of assimilating the conditions of related problems for the sake of their resolution.

It is also worth observing that the success of a formal method may be conspicuous without thereby being absolute. Especially adverse conditions are capable of rendering an otherwise fruitful approach inapplicable. If a circuit is not available around $\mathcal{G}$, we have seen that one can be designed in an optimal manner. However, even if a circuit can be designed, it may not determine an optimal sweeping schedule. If it takes longer than the duration of relevant parking restrictions to travel

---

[14] The degree of a node is the difference between the number of incoming edges and the number of outgoing ones.

[15] Given a street-network like $\mathcal{E}$, in which every two nodes are linked by a sequence of directed edges, and two nodes u, v from $\mathcal{E}$, Dijkstra's algorithm is an effective method to determine the shortest path from u to v (recall that edges in $\mathcal{E}$ are weighted by travel times). In a nutshell, the algorithm is initiated by assigning 0 to u and ∞ to every other node in $\mathcal{E}$. After visiting u, the algorithm recursively visits the unvisited node with the shortest known distance from u. Updated numerical labels are assigned to nodes at each iteration of the algorithm: whenever the distance from u to some node z is shorter than the distance previously assigned to z, an update occurs. Distances from u are progressively updated until every node is visited. When the algorithm stops, the shortest distance of u from every other node is known, and, in particular, its shortest distance from v is. The algorithm was originally presented as Problem 2 in Dijkstra (1959).

[16] For details, see Tucker and Bodin (1976, pp. 85–84, 92–95).

around the circuit, more than one mechanical broom must be employed to sweep the same area. The full Eulerian circuit has to be broken into shorter, feasible paths, each traversed by a distinct mechanical broom. It is possible that the whole tour may be optimal,[17] while its break-up into sub-routes is not. In such a case, Eulerian routing is not of help. Routing done by hand must then replace it to determine variants of the overall optimal route that can be broken into fewer subroutes.[18]

It is now easy to see that the street sweeping problem just discussed displays a range of characteristic interactions between mathematical methods and empirical conditions. The latter appear as more or less severe adversities, in the face of which mathematical methods have to be enriched and adapted: adversities promote the construction of mathematical strategies but may also prove definitive stumbling blocks for a given methodology. In this respect, problem-solving activities appear mainly as a reckoning with challenging circumstances. It is clear, however, that formal methods do not merely respond to empirical conditions: they also structure and direct enquiry. More precisely, they can codify the terms of a problem in such a way that its resolution becomes amenable to a chosen mathematical approach. A relatively recent instance of this phenomenon, which involves Eulerian routing, has arisen in genetics.

4 Sequencing

While discussing optimal street sweeping schedules, I focussed on a problem-solving context in which a given method was progressively articulated to confront a range of possible empirical adversities. The reverse of this phenomenon occurs, too. The terms of a given problem may be structured in such a way that they make a prescribed method applicable. In this case, instead of responding to empirical adversities, a problem-solving strategy creates an empirical opportunity. Usually, the organisation of data that enables a specific formal treatment is simultaneous with a variation or modulation of the treatment itself.

An application of Eulerian routing to genetic sequencing provides an illuminating illustration of what has just been described in general terms. A central problem of genetic sequencing is called shotgun fragment assembly: it arises from the fact that complete DNA strands cannot be read continuously. For this reason, whenever a genome is being mapped, longer sequences of nucleotides are broken (e.g. by ultra-sound bursts) into shorter fragments that can be read and, for this reason, are known as 'reads'. The goal is then to assemble the original sequence from the fragmentary reads, thus reconstructing the original genome. A standard technique to achieve this goal is known as the 'overlap-layout-consensus' approach.

---

[17] In the sense that it takes the shortest time and minimises inconvenient turns as well as deadheading time.

[18] An example formally involving a disconnected, directed graph, whose components have to be linked in an overall tour, is discussed in Tucker and Bodin (1976, pp. 99–100).

The approach is based codifying reads as nodes and overlaps between reads as directed edges. The resulting configuration $\mathcal{G}$, known as an overlap graph, can be used to tackle the assembly problem: the assembled sequence may be determined by finding a path that visits each node exactly once, also known as a Hamiltonian path.[19] Besides the problem of correcting errors after a path has been determined, this problem-solving strategy is seriously affected by the fact that no efficient algorithm is available to find Hamiltonian paths on an arbitrary, finite graph (if $P \neq NP$, none exists).

A further problem is caused by the fact that certain genomes exhibit many long, identical sequences, known as repeats. Thus, if R is a repeat occurring three times and A, B are reads whose ends overlap R, the available information does not immediately suggest whether the correct assembly is *RARBR* or instead *RBRAR*. The problem of determining the right order of reads in presence of repeats is known as the 'repeat' problem.

It is largely for the purposes of finding an effective algorithm to tackle the repeat problem that Pevzner et al. (2001) introduced an approach to fragment assembly based on Eulerian routing. The immediate appeal of an Eulerian approach is due to the fact that Eulerian routes, if they exist, can be determined in linear time.[20] Since the original approach to fragment assembly is already stated as a path-finding problem, it is plausible to conjecture that Eulerian routing might be applicable.

The issue of greater interest, for present purposes, is that the assembly problem is not standardly formatted as an Eulerian routing problem. To apply Eulerian routing to this problem, it is necessary to restructure reads and overlaps to a suitable specification. Given a set S of reads, the key idea is to look at S_l, the set of all sub-reads of length l, also known as l-mers (for example, if $l = 7$, the string of nucleotides CGT GCAA is a l-mer). The items in S_l are structured as a configuration known by the name of a De Bruijn graph. A De Bruijn graph is obtained by (i) regarding $(l-1)$-mers as nodes and (ii) assigning a directed edge from node u to node v if, and only if, there is $s \in S_l$ such that u describes the first $l-1$ positions of s and v describes the last $l-1$ positions of s.

Eulerian routing is a meaningful strategy on a De Bruijn graph. To see why, consider the artificially simple case[21] of a circular genome and suppose that S_l contains every possible l-mer. Then every $(l-1)$-mer must be both the prefix of a l-mer and the suffix of some other l-mer. In other words, the incoming and outgoing edges balance out at every node and the S_l can be assembled into an Eulerian circuit.

In general, the assembly problem is significantly more complicated than the toy example just described. One key empirical adversity is posed by repeats. If copies of the same, repeated l-mers are identified, then the routing goal cannot be an Eulerian path, because some repeated sequences may have to be traversed multiple times. It follows that, if Eulerian routing is to be applied, repeated l-mers should not be

---

[19] The actual procedure is much more complicated, but what matters here is its 'routing' content.

[20] If V is the number of nodes and E the number of edges of a given graph, Hierholzer's algorithm decides whether an Eulerian route exists in a time that is a fixed constant multiple of $V + E$.

[21] Taken from Compeau et al. (2011, p. 989).

identified but instead treated as distinct edges. A correct assembly now corresponds to an Eulerian route that goes through multiple edges associated with the same *l*-mer *in the right order*. The issue is therefore to find an Eulerian path that contains specific subpaths.[22] Piecing subpaths together into a trail is what Pevzner et al. (2001) calls an *Eulerian Superpath Problem* (ESP).

The ESP may be regarded as a modulation of Eulerian routing because it can be turned into an Eulerian routing problem only by means of a suitable reduction technique. The situation is qualitatively different from the street sweeping problem examined in the previous section: in that case, empirical adversities required the introduction of already available algorithms and their integration into a broader problem-solving strategy. In the present situation, the adversity originated by the repeat problem requires a new technique under which ESP may be regarded as a variation of Eulerian routing.

To reduce ESP to Eulerian routing, one starts from an extended De Bruijn graph $\mathcal{G}$ with new 'repeated' edges and its associated set of paths $\mathcal{P}$. The pair $(\mathcal{G}, \mathcal{P})$ is then reduced to a pair $(\mathcal{G}_1, \mathcal{P}_1)$ in such a way that: (i) $\mathcal{G}_1$ has fewer edges than $\mathcal{G}$; (ii) there is a function $f : \mathcal{P} \longrightarrow \mathcal{P}_1$ whose restriction f_1 to Eulerian Superpaths is a one-to-one correspondence.

Repeated reductions are carried out until a finite sequence of the form:

$$(\mathcal{G}, \mathcal{P}) \longrightarrow \dots \longrightarrow (\mathcal{G}_n, \mathcal{P}_n) \tag{1}$$

results, where P_n is an edge of $\mathcal{G}_n$. The information encoded by the pair $(\mathcal{G}_n, \mathcal{P}_n)$ and the bijections $f_i (i = 1, \dots, n)$ makes it possible to use Eulerian paths in $\mathcal{G}_n$ to find Eulerian Superpaths in $\mathcal{G}$. The path-complexity reduction described in (1) is or is not viable depending on the available data.

To see this, note that, in order to reach a set of edges in $\mathcal{P}_n$, it is necessary to delete edges along the way. An obvious type of deletion consists in eliminating one of two consecutive edges: this operation is called detachment. Detachment is unproblematic if there is exactly one directed path through the given edges. It then reduces the length of at least one path by one edge. Suppose, for instance, that the directed path *CR*, with *C* preceding the repeat *R*, is under consideration, and that the available reads provide the data *DRB* and *CRA*. In this case, it is still possible to detach *CR*, linking the vertex from which *C* emanates to the vertex from which *A* emanates. The direction *CA* is fixed by the available reads. The path *CR* is in this case said to be *resolvable*, because it can be included in a path constrained by the data.

Other cases in which the repeat *R* is consecutive with at least two incoming reads or two outgoing reads pose problems. For instance, the information *CRB* and *CRA* requires, but does not provide, an order of priority. Some paths may not be resolvable, e.g. because they are compatible with several extensions. In this case the

---

[22] The relevant subpaths are obtained from the De Bruijn graph by taking several distinct copies of edges that codify repeats. The subpaths of interest are thus not subpaths in the De Bruijn graph but paths determined by that graph, to be eventually included as subpaths into the Eulerian path whose construction is described below.

analysis must continue[23] and transformations other than detachment must be introduced. Paths may also be incompatible with all the given extensions, in which case the ESP task cannot be completed. The available data and the results of previous stages in the ESP reduction process determine what subsequent actions are possible. We see, again, an epistemic activity that progresses through stages, responding to empirical constraints with an array of formal moves, which describe a modulation of Eulerian routing called for by a specific interest in superpaths, as opposed to simple paths.

The great advantage of working with this modulation of Eulerian routing is computational: the approach is not as time expensive as its Hamiltonian counterpart. It is worth noting that this fact alone does not completely undermine the earlier Hamiltonian approach to genetic sequencing. It is known that this approach is computationally expensive and becomes unserviceable when standard algorithms for large Hamiltonian path problems are run on a silicon-based computer. Over the last few decades, however, it has been suggested that Hamiltonian path problems of an otherwise forbidding size might be successfully tackled on a DNA computer.[24]

This suggestion, as well as the research surrounding it, deserves some attention because it sheds light on the impact of technology on mathematical problem-solving. Until now, my major focus has been the interplay between problems, mathematical methods and goals within scientific enquiry. These three factors enter a fourfold interplay that involves technological capabilities and their variation over time. The next section is devoted to exploring the impact of technology on mathematical problem-solving with reference to the Hamiltonian path problem. I shall not turn to this issue immediately, however, but only after situating it in a context that can be clearly outlined only by reflecting further on the what has already been said about problem-solving.

5 A Fourfold Interplay

Sections 3 and 4 highlighted interesting interactions between three factors: empirical problems, the goals these problems set to scientific enquiry and the formal methods deployed to achieve them. A particularly important phase of the interaction between these factors has been noted by Otávio Bueno and Mark Colyvan when they wrote that 'the world does not come equipped with a set of objects [...] and sets of relations on those' (Bueno & Colyvan, 2011, p. 347).

It is now possible to refine their observation and set it within a wider theoretical context. Because scientific enquiry may be regarded as a complex of problem-solving activities, the goals set by given problems and the methods introduced to tackle

---

[23] See Pevzner et al. (2001) for 9752–9753 for details omitted here.

[24] The focus on the Hamiltonian path problem does not arise from the requirements of genetic sequencing. Because this problem is NP-complete, i.e. every other problem in the same time-complexity class is polynomially reducible to it: it follows that any convenient method to tackle the problem transfers to every problem in the same class. Hundreds are known. For a list, see e.g. Garey and Johnson (1979, p. 180) and ff.

them guide the structural specification that is eventually selected for the terms of any given problem, i.e. the available empirical information.[25] Thus, for instance, the same problem may lead to distinct structural specifications, required by alternative methods aimed at solving it. Furthermore, the same empirical setting is subjected to distinct formal treatments when it provides the underlying reference for distinct problems.

Section 4 illustrated the impact of method selection on the way the terms of a problem are structured. The earlier discussion of genetic sequencing showed that the De Bruijn graph is constructed by 'reversing' the role of relata (vertices) and relations (edges) in the overlap graph. The latter graph codifies overlaps by edges and reads by vertices, whereas the former codifies overlaps (of a fixed length) by vertices and reads (of a fixed length) by edges. The data supplied by the assembly problem are just k-mers, exhibiting no differentiation in type. A differentiation emerges only after the selection of a problem-solving method has been effected. If Eulerian routing, suitably modulated, is adopted, then certain strings of nucleotides are treated as nodes and others as edges. If Hamiltonian routing is adopted instead, a dual selection is carried out: strings treated as edges by the Eulerian methodology are now treated as nodes, and vice versa.

Section 3 does not naturally lend itself to an illustration of the last point[26] but indirectly highlights the fact that the same type of empirical information may be subjected to very different treatments, depending on the goal at hand. For instance, a street network can be the background against which distinct tasks, e.g. street sweeping or post delivery, must be carried out. With respect to street sweeping, the network may be treated as shown in Sect. 3. With respect to post delivery, the network is treated in Hamiltonian terms: more precisely, the goal of an optimal delivery programme is to find the shortest route that visits each customer exactly once. Achieving this goal is the same as solving what is known as the shortest Hamiltonian path problem (sHPP).[27]

Against the background of problem-solving one may therefore see a constant interplay of empirical information, formal methodologies and goals. The selection of goals may affect the choice of formal methods and the selection of methods may determine the structural specification to be employed. This picture is enriched by one additional factor, namely the impact of technological advances on the viability of formal methods. The fourfold interplay between problems, goals, methods and

---

[25] As observed in Sect. 2, the construction of a method also guides the selection of information.

[26] It would not be reasonable to design a computationally intractable alternative to an existing Eulerian routing approach.

[27] It is worth noting, if incidentally, that the problem itself is in general modulated on the conditions of operation: if no Hamiltonian path exists, one may seek to solve a shortest pre-Hamiltonian path problem, which only requires that each customer should be visited at least once. If not all customers within a district are visited on the same day, a variation of the shortest pre-Hamiltonian path problem, known as the shortest pre-Hamiltonian rural path problem, becomes pertinent.

technological advances shapes the context of enquiry in problem-solving by mathematical means.[28]

In order to clarify the role and effects of technological advances, it is, at this point, quite natural and convenient to look at genetic sequencing again: whilst graph-theoretical methods have been adopted to tackle the assembly problem, as discussed in Sect. 4, the possibility of codifying graphs by means of k-mers has been exploited to design DNA computers (working in vitro or in vivo) that can solve HPP (for sufficiently small graphs, to date). This is a phenomenon of special interest because, rather than conforming to the pattern witnessed so far, in which mathematical resources codify empirical information, it reverses it: molecular resources now codify graph-theoretical information.

Work with a DNA computer was pioneered in Adleman (1994), whose approach still provides a useful template for current research [see e.g. Sergeenko et al. (2020)]. On Adleman's approach, a graph is given and k-mers are engineered to codify its vertices and edges. In order to understand how the coding works, it is necessary quickly to recall some basic facts concerning the structure of DNA. The DNA is a molecule consisting of two strands linked by hydrogen bonds. Each strand is a sequence of nucleotides, macromolecules containing a pentose sugar known as 2'-deoxyribose, to which a base (A, T, C or G) and a phosphate group (PO_4) are attached. The five carbon atoms in the sugar molecule are referred to as i', with $i = 1, 2, 3, 4, 5$. A k-mer is thus a string of nucleotides joined together by phosphate groups linking the 3' carbon of one sugar to the 5' carbon of the next sugar. The direction $5' \rightarrow 3'$ is the direction of DNA replication.

Adleman codified each vertex of a seven-vertex digraph by means of 20-mers. In his codification, a directed edge from vertex u to vertex v in the graph corresponds to the 20-mer obtained by splicing the final 10-mer u and the initial 10-mer of v, read in the $5' \rightarrow 3'$ direction. If the start x and the end y of a desired Hamiltonian path are fixed, then any edge emanating from x is codified by the whole 20-mer x followed by the initial 10-mer of the vertex adjacent to it. In a similar manner, edges on which y is incident are codified by a 30-mer including all of y.[29]

Once a digraph, and more specifically the set of its edges, is encoded by oligonucleotides, it is possible to make them ligate in order to produce random sequences of edges (which may not be Hamiltonian paths or even directed paths). The advantage of this procedure is that it is relatively easy to produce large numbers of copies of the same edge (Adleman worked with 3×10^{13} copies per edge) and thus generate, in a one-shot fashion, a very large number of edge sequences. Their random generation implements the first step of a nondeterministic HPP algorithm. The subsequent steps of the algorithm consist in selecting the randomly generated paths starting at x and ending at y (Step 2). Next, one selects, among these paths, the ones visiting exactly n vertices (Step 3). Finally, the paths visiting each distinct vertex at least once are

---

[28] Thus Bueno and Colyvan have indirectly isolated one moment (structuring empirical information for the sake of specific intervention) within this broader interplay.

[29] Note that, in this codification, the polarity of *DNA* strands is exploited to discriminate the orientation of edges in a digraph.

singled out from those isolated in Step 3 (Step 4). If any exist, a Hamiltonian path from *x* to *y* has been found. If none exists, no Hamiltonian path exists (Step 5).[30]

Without going into too many technical details (Adleman, 1994, p. 1022) offers a very readable account), it is worth pointing out that Steps 1, 2 and 4 rely on a decisive use of Watson–Crick complementarity. It is well known that the only base pairings in the DNA molecule are A-T and G-C (in either order). Given a sequence of bases, its Watson-Crick complementary is obtained by turning *A* into *T*, *T* into *A* and doing the same with *G*, *C*. A 20-mer *u* on a DNA strand will thus combine with its complementary *ū*. In Step 4, the complements of edges are used to check that paths visit each vertex exactly once (e.g. if *u* is visited, then the strand containing it combines with *ū*). In Step 2, *x* and *ȳ* are used to initiate DNA replication in order to single out edge sequences starting with *x* and ending with *y*. In step 1, complements of vertices are used as splints to support ligations.

The procedure just described affords massive parallel processing at the initial step: as I have already noted, a very large number of edge-sequences, i.e. attempts at producing a solution, can be generated at once. Moreover, as remarked in Sergeenko et al. (2020, p. 73), only Step 4 is time expensive and its time consumption is linear in the number of vertices. By contrast, silicon-based computers that run standard HPP algorithms give rise to exponential time consumption when a graph is large and does not have many edges.[31] This is why DNA computing holds great promise.[32]

For present purposes, its relevance is twofold. First, it shows that the relation between formal features and empirical traits in problem-solving is, to some extent, reversible. In Sects. 3 and 4, Eulerian routing and its variations arose as a methodology subjecting empirical information to formal treatment. In DNA computing, quite the opposite is the case: the formal configuration of a graph is subjected to molecular encoding for the sake of implementing a computation. Second, the motivation and prospects of DNA computing suggest that the concrete design of a problem-solving strategy may regulate the conditions of its applicability: algorithms that are unserviceable when run on silicon-based computers to tackle large problems could regain their usefulness under a different implementation.

I noted earlier that empirical adversities may stimulate the articulation of a mathematical problem-solving methodology. In view of the last observations, it is also possible to observe that an existing mathematical methodology can be favoured by empirical opportunities. This interrelationship between empirical conditions and mathematical methods underlies mathematical problem-solving in scientific practice. Eulerian routing has provided a fruitful starting point to offer a concrete illustration of the subtle manner in which problems setting goals and calling for methods open the way to developments and modulations of formal methodologies, which

---

[30] Provided one has generated Hamiltonian paths in the initial random step.

[31] In technical terms, when its incidence matrix is sufficiently sparse. See Sergeenko et al. (2020, p. 70) for a discussion. The estimates of the energy efficiency and computational power associated with DNA computing in Adleman (1994, p. 1023) are also of interest in this connection.

[32] For discussion of issues attending its implementation, see e.g. Adleman (1994, p. 1023), Baumgarden et al. (2009, p. 8) and Sergeenko et al. (2020, p. 73).

unfold under the sway of empirical conditions. The emerging picture of problem-solving by mathematical means does not only shed light on the structure of scientific practice. It can also be used productively to reconsider many insights from the philosophical literature on mathematical explanation. I am going to show how in the next section.

6 Reconstructing Explanatory Analyses

The paradigmatic, non-trivial instances of mathematical explanation encountered in the philosophical literature describe especially well-behaved settings. These settings are static in the sense that they do not involve ongoing eqnuiries but settled facts and, since enquiry is no longer unfolding, they are not sites for the elaboration of problem-solving methods but stages on which explanatory accounts transparently appear. Although it is reasonable to pay selective attention to explanatory circumstances, on account of their distinctive features, it is important not to lose sight of their genealogy. The settled facts accounted for by a mathematical explanation are available because data were selected and inserted into an inferential trajectory leading to the construction of methods that could make the data function as clues to set outcomes. At the close of enquiry, formal methods are left, together with a clear understanding of the conditions and the outcomes they link. Because it becomes possible to set up such links deliberately, explanatory opportunities arise.

The last remarks offer a productive way of looking at the insights into mathematical explanation provided in philosophical work. If these insights arise from restricting attention to contexts that depend on an antecedent set of epistemic activities, it may well be possible that their significance is not confined to the explanatory event alone, but may be ascribed to the whole trajectory of scientific-enquiry. There is no reason, it seems to me, to want to force wide-ranging insights into the Procrustean bed of mathematical explanation. There are, in fact, at least two reasons not to do it. First, salient aspects of the application of mathematics that can be identified under explanatory circumstances are more richly portrayed when they can be ascribed to problem-solving endeavours, as opposed to some of their possible terminations. Second, the static context of explanation, in which facts are settled and formal results neatly account for them, lends itself to metaphysical formulations that become unnecessary as soon as it is possible to refer explanatory results back to the epistemic activities on which they ultimately depend.

These observations are quite general and their actual content may yet prove difficult to evaluate. In order to clarify them, I now turn to a few, insightful analyses of mathematical explanation encountered in the work of Alan Baker and Christopher Pincock, who devoted much effort to this topic. My goal is to show that, once Baker's and Pincock's analyses are referred to the context of enquiry, they achieve greater articulation and sophistication. Moreover, they can be liberated from metaphysical connotations naturally suggested by a restriction to the circumstances of explanation, insofar as they are divorced from the epistemic activities they presuppose.

I turn first to Christopher Pincock's conception of abstract explanation [put forward in Pincock (2007) and refined in Pincock (2015a)[33]). Pincock isolates this species of explanation with reference to Euler's routing criterion, which, in Pincock (2007), he regards as an abstract explanation of the solved Königsberg bridge problem. In essence, Pincock observes that the non-traversability of the Königsberg bridges along a trail is transparently accounted for by the formal features of an abstract graph that can be associated with a geographical configuration. In Pincock (2015a), the notion of abstract explanation is qualified in general terms while being compared with the distinct notion of a programme explanation, which I won't discuss here [see Lyon (2012) for an application of this notion to mathematical explanation]. Pincock notes that:

> Programme explanations and abstract explanations both appeal to what is more abstract than the phenomenon being explained. However, abstract explanations invoke a more abstract entity and its properties. Programme explanations appeal only to a more abstract property of the physical system itself. This might not seem like such a big difference, but it has important implications for the features that are central to the explanatory value of abstract explanations. We get necessary and sufficient conditions for the explained property to apply as well as an informative comparison between novel kinds of objects. (Pincock, 2015a, 2015b, p. 873)

As transparently evinced by the use of diagrams in Pincock (2007, p. 258), the Königsberg bridge configuration can be involved in an abstract explanation because of its correspondence with an abstract graph and its properties. The graph's edges stand for connecting bridges and its nodes for land masses. An explanatory relation is thus established between a geographical configuration and a diagram or perhaps, given Pincock's insistence on abstractness, on the isomorphism type of a finite graph. It seems to me that Pincock's conception of an abstract explanation is introduced because it allows him sharply to identify the terms of a dependence relation, i.e. a specific, abstract graph and a concrete configuration. Such a dependence relation is what enables him explicitly to spell out a role for mathematical resources, here thought of as mathematical objects, in applications.

It seems to me that Pincock is right to emphasise the dependence of a routing impossibility upon combinatorial considerations of a graph-theoretic nature, but I think that his notion of dependence would acquire even greater significance if it could be ascribed to the whole trajectory of enquiry whose close has among its byproducts a simple and elegant remark about the non-traversability of the Königsberg configuration. In fact, if attention is moved away from this latter fact to encompass the work done by Euler to establish a routing criterion, it is easily seen that Euler never makes special use of the Königsberg bridge configuration (he only works with finite strings of letters), whilst constantly trying to articulate the

---

[33] Pincock relies upon the same conception in his discussion of a classic application of Galois theory to the unsolvability of the general quintic [see Pincock (2015b)]. Even though this example is not relevant to empirical science, a version of the considerations to follow can apply here.

dependence of a solution to the routing problem upon this problem's terms. The inferential trajectory, discussed in Sect. 2 in detail, that leads Euler to a routing criterion provides a way of interacting with contents that can be run on a variety of different artefacts by a variety of agents. It can be run equally well by a student looking at the a graph-theoretic diagram and a surveyor looking at a system of bridges from the top of a hill.

The pervasive form of dependence at work here is of an inferential kind: it can be ascribed to a formal method because the latter method is capable of selecting and using data as pointers to a problem's resolution. When, however, the construction of a method is put aside and a rigorous restriction to special explanatory circumstances is demanded, the only items available are a graph-theoretic argument attached to a special configuration and a settled fact. In this context, it is natural to try and connect these localised features of a broader and antecedent process to objects (an empirical setting, an abstract mathematical object). The connection, however, is largely a result of the decision artificially to insulate explanatory outcomes from the enquiries that make them possible.

A quick look at the modern approach to Euler's routing criterion confirms the last remarks. Because the goal is to determine necessary and sufficient conditions for the existence of a trail in a graph, it is not helpful to focus on special abstract configurations in the first instance (each of them is a special case and, thus, exhibits features possibly irrelevant to a general existence criterion). A combinatorial approach succeeds when it can rise to the description of a type of interaction with a generic configuration. A key feature of the interaction of interest requires that, if an intermediate locality is reached through an edge x_1, it must be left through an edge $x_2 \neq x_1$. Thus, every intermediate locality must be incident on an even number of edges. These remarks are not reserved for abstract configurations: if they were, graph-theory would be inapplicable. They nonetheless establish the dependence of the resolution of a problem upon a computable feature of its terms.

Pincock's choice to characterise abstract explanations in terms of dependence is a felicitous one. His restriction to the circumstances of explanation alone, however, unduly restricts the scope of his analysis and unhelpfully regiments its formulation. I have tried to show that such restriction removes from sight the fact that explanatory accounts of the kind Pincock favours can only be produced against the availability of a method that crystallises effective inferential paths in problem-solving. Once this is recognised, the notion of dependence Pincock insists upon can be reconstructed as the inferential dependence, mediated by mathematical method, of the solution to a problem upon the problem's terms. This kind of dependence is frequently encountered and does not mark explanatory settings alone since it is of central importance to problem-solving as a whole.

Similar considerations apply to the important analysis of generality found in Baker (2017), which can be related to interesting antecedents in Baker (2005, 2009). Baker repeatedly stresses the fact that mathematical resources are significant in explanatory contexts on account of their generality and has provided interesting specifications of these notions, which are closely examined in Baker (2017). In the

latter paper, as well as his earlier work,[34] Baker discusses generality as a quality of the entities that enter mathematical explanations because generality cannot be referred to empirical particulars. The significance of generality, when ascribed to structured mathematical objects, is that it identifies their role as schemas or patterns under which empirical particulars can be organised or, more precisely, find themselves organised, since all facts are settled when only a mathematical explanation is required. Abstract schemas are both *scope general*, in the sense that they apply to many homogeneous empirical configurations, and *topic general*, because they can be transferred across heterogeneous empirical configurations.

These notions of generality, which Baker ascribes to mathematical entities under explanatory circumstances, are naturally exemplified and ascribed to mathematical methods in the non-explanatory contexts from Sects. 3 and 4. Eulerian routing is a scope general method because it can be used on city districts, irrespective of their layouts. Different layouts are homogeneous in that they convey the same type of empirical information relative to a goal of reference like street sweeping. Eulerian routing is also topic general because its applicability is not bound to empirical information of the kind displayed by street networks. Section 4 showed that Eulerian routing can be transferred to the rather different domain of genetic sequencing, which provides correspondingly different empirical information. These remarks already clarify that Baker's notion of generality does not have to be tied to the circumstances of explanation or assigned to mathematical entities, because mathematical methods are deliberately framed to achieve at least scope generality and problems may be framed in such a way that they enable formal methods to acquire topic generality.

Moreover, the ascription of generality to mathematical entities, which is again suggested by the static character of explanatory contexts, prevents Baker from seeing his notion as a special case of mathematical plasticity, namely the adaptability of inferential trajectories to varying circumstances and varying problems. Under explanatory circumstances, one must have a structure that exactly accounts for a given fact. The ordinary circumstances of scientific enquiry do not exhibit such a tight correspondence but, as I have repeatedly pointed out in the preceding sections, present more or less severe empirical adversities that require the articulation and modulation of mathematical methods. In other words, methods may be immediately scope or topic general when the circumstances are fortunate enough to guarantee the possibility of lifting them from one context of enquiry to another. In general, empirical circumstances are not so generous but methods can be transferred nonetheless, provided they are modified and adapted. Thus, for instance, Eulerian routing does not solve the general street sweeping problem without being integrated into a more comprehensive and complicated problem-solving strategy.

In a similar manner, Eulerian routing does not tackle genetic sequencing unless it is modulated in response to the repeat problem, which leads to a search for superpaths. These are typical forms of plasticity. When articulation and modulation are not required, one obtains Baker's generality. The notion is therefore important but

---

[34] See especially Baker (2009, p. 617) and Baker (2005, pp. 224–233).

its explanatory formulation makes it difficult to discern that generality is a special case of plasticity precisely because the stress on mathematical entities (e.g. abstract graphs) emphasises structural rigidity, as opposed to methodological plasticity. In short, once Baker's notion of generality is freed from a rigorous explanatory connotation, its broader significance becomes easier to appreciate and can be naturally disengaged from a metaphysical formulation that does not help appreciate the versatility of mathematical treatment.

It seems to me that other discussions of mathematical explanation can be subjected to the reconstruction I have outlined with reference to Pincock's and Baker's work.[35] The advantage in doing so is, as I tried to show, twofold. On the one hand, important insights can be properly ascribed to the whole pathway of problem-solving enquiry, as opposed to its possible explanatory *coda*. On the other hand, the methodological content of these insights becomes more transparent when freed from metaphysical formulations that seem natural only with reference to the static explanatory context.

7 Concluding Remarks

Mathematical explanation, as it has been discussed in the recent philosophical literature, has a distinctive synchronic quality. A fact and a mathematical result available to account for it are taken as given at once and their relation is thematised as a philosophical problem. This synchronic perspective, however prominent, has not been strictly adhered to. For instance, Baker's work on topic generality draws attention to the fact that explanatory resources, once available, may admit of transfer to new explanatory contexts at a later time. As a further example, Pincock's work on the mathematical account of Plateau's laws in Pincock (2015a) follows the development of mathematical methods geared towards the analytical resolution of abstract optimisation problems [in Pincock (2015a)]. In both cases, a tendency emerges towards looking at a research context that goes beyond a single, self-contained explanatory moment. Both Baker and Pincock have thus recognised the pertinence of a diachronic perspective within their work on explanation.

In this paper, I have tried to turn their recognition into an explicit methodological orientation, using as my leading example an application of mathematics frequently discussed in explanatory terms, namely the Königsberg bridge problem. Moving away from its explanatory framing, I sought to examine both the epistemic activities that precede this problem's resolution and the epistemic activities that drive its articulation and modulation in later enquiries. My principal aim was to show that a shift of focus from explanatory concerns to mathematical problem-solving in the context of scientific enquiry can be philosophically very rewarding because it leads to a rich and comprehensive picture of the application of mathematics.

---

[35] For instance, the account of programming explanations in Lyon (2012) is amenable to reconstruction, on account of its proximity to the notion of abstract explanation discussed above.

While significant in itself, the resulting picture can also help broaden and sharpen insightful analyses that have been originally formulated in connection with mathematical explanation. Such analyses are broadened because they no longer need to be restricted to explanatory contexts but can be referred to the whole trajectory of enquiry. The analyses are also sharpened because their metaphysical formulation, which suggests itself as natural once the context of enquiry is suppressed, tends to conceal their actual character and import.

References

Adleman, L. (1994). Molecular computation of solutions to combinatorial problems. *Science, 266*, 1021–1024.

Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind, 114*, 223–238.

Baker, A. (2009). Mathematical explanation in science. *British Journal for the Philosophy of Science, 60*, 611–633.

Baker, A. (2017). Mathematics and explanatory generality. *Philosophia Mathematica, 25*, 194–209.

Baumgarden, J., Acker, K., Adefuye, O., Crowley, S. T., DeLoache, W., Dickson, J. O., et al. (2009). Solving a Hamiltonian path problem with a bacterial computer. *Journal of Biological Engineering, 3*, 11.

Bodin, L. D., & Kursch, S. J. (1978). A computer-assisted system for the routing and scheduling of street sweepers. *Operations Research, 26*, 525–537.

Bueno, O., & Colyvan, M. (2011). An inferential conception of the application of mathematics. *Noûs, 45*, 345–374.

Chang, H. (2011). The philosophical grammar of scientific practice. *International Studies in Philosophy of Science, 25*, 205–221.

Compeau, P. E. C., Pevzner, P. A., & Tesler, G. (2011). Why are de Brujin graphs useful for genome assembly? *Nature Biotechnology, 29*, 987–991.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 1*, 269–271.

Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae, 8*, 128–140.

Fleischner, H. (1990). *Eulerian graphs and related topics (part 1, volume 1).* Annals of mathematics 45. North-Holland.

Gaifman, H. (2005). Some thoughts and a proposal in the philosophy of mathematics. In P. H. Hajek, L. Valdes-Villanueva, & D. Westerstahl (Eds.), *Proceedings of the twelfth international congress logic, methodology and philosophy of science* (pp. 351–372).

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. W.H. Freeman and Company.

Jansson, L., & Saatsi, J. (2019). Explanatory abstractions. *The British Journal for the Philosophy of Science, 70*, 817–844.

Kuorikoski, J. (2021). There are no mathematical explanations. *Philosophy of Science, 88*, 189–212.

Lange, M. (2013). What makes a scientific explanation distinctively mathematical? *The British Journal for the Philosophy of Science, 64*, 485–511.

Lyon, A. (2012). Mathematical explanation of empirical facts, and mathematical realism. *Australasian Journal of Philosophy, 90*, 559–578.

Pevzner, P. A., Tang, H., & Waterman, M. S. (2001). An Eulerian path approach to DNA fragment assembly. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 9748–9753.

Pincock, C. (2007). A role for mathematics in the physical sciences. *Noûs, 42*, 253–275.

Pincock, C. (2015a). Abstract explanations in science. *The British Journal for the Philosophy of Science, 66*, 857–882.

Pincock, C. (2015b). The unsolvability of the quintic: A case study in abstract mathematical explanation. *Philosopher's Imprint, 15*, 1–19.

Räz, T. (2018). Euler's Königsberg: The explanatory power of mathematics. *European Journal of Philosophy of Science, 8*, 331–346.

Sergeenko, A., Yakunina, M., & Granichin, O. (2020). Hamiltonian path problem solution using DNA computing. *Cybernetics and Physics, 9*, 69–74.

Tucker, A. C., & Bodin, L. D. (1976). A model for municipal street sweeping operations. In F. S. Roberts (Ed.), *Discrete mathematics with applications to social, biological and environmental problems* (pp. 76–111). Prentice-Hall.