AppliedMath

MDPI

Tutorial

On the Differential Topology of Expressivity of Parameterized Quantum Circuits

Johanna Barzen * and Frank Leymann *

Institute of Architecture of Application Systems (IAAS), University of Stuttgart, Universitätsstr. 38, 70569 Stuttgart, Germany

* Correspondence: johanna.barzen@iaas.uni-stuttgart.de (J.B.); frank.leymann@iaas.uni-stuttgart.de (F.L.)

Abstract

Parameterized quantum circuits play a key role in quantum computing. Measuring the suitability of such a circuit for solving a class of problems is needed. One such promising measure is the expressivity of a circuit, which is defined in two main variants. The variant in focus of this contribution is the so-called dimensional expressivity, which measures the dimension of the submanifold of states produced by the circuit. Understanding this measure needs a lot of background from differential topology, which makes it hard to comprehend. In this article, we provide this background in a vivid as well as pedagogical manner. Especially, it strives towards being self-contained for understanding expressivity, e.g., the required mathematical foundations are provided, and examples are given. Also, the literature makes several statements about expressivity, the proofs of which are omitted or only indicated. In this article, we give proof for key statements from dimensional expressivity, sometimes revealing limits for generalizing them, and sketching how to proceed in practice to determine this measure.

Keywords: quantum computing; differential topology; parameterized quantum circuits; expressivity; variational quantum algorithms

check for updates

Received: 24 July 2025
Revised: 26 August 2025
Accepted: 1 September 2025
Published: 4 September 2025

Citation: Barzen, J.; Leymann, F. On the Differential Topology of Expressivity of Parameterized Quantum Circuits. *AppliedMath* **2025**, *5*, 121. https://doi.org/10.3390/appliedmath5030121

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Quantum algorithms are unitary operations transforming an initial state into another state that either directly represents the solution or that is a state that is measured, and the measurement result leads to the solution of the problem [1]. Quantum algorithms are typically realized as quantum circuits: a composition of unitary operations at lower dimensions that represent the decomposition of the overall unitary transformation. These lower-dimensional unitary operations are typically the operations supported by a concrete quantum computer. Sometimes, the unitary operations are parameterized: for example, these parameters are angles by which a state must be rotated. By varying the parameters between different executions of the circuit, a solution is iteratively computed.

Such variational quantum algorithms (VQAs) [2] and especially their encompassed parameterized quantum circuits (PQC) (Section 2.1) gain a lot of attention not only because of their appropriateness for today's available computers (which are noisy [3]) but also because parameterized quantum circuits represent applications in themselves (like quantum neural networks [4,5]).

Whether or not a variational quantum algorithm can compute a solution of the problem addressed depends on the "suitability" of the parameterized quantum circuit it encompasses. Thus, having a measure of "suitability" in this context is of utmost importance.

Such a measure is called *expressivity* of the parameterized quantum circuit and has been defined in two major variants:

- First, a parameterized quantum circuit associates with each parameter tuple a unitary matrix. The more of the unitary group is covered by the parameterized quantum circuit, the better. This is because if circuits that solve the problem (i.e., unitary matrices) exist, then the unitary matrices generated by the VQA should include them. And the likelihood that this is achieved is higher the more of the unitary group is covered. The coverage of the unitary group is suggested to be assessed by means of how close the generated matrices are to being Haar random [6]. This is what we call the "unitary approach"; it is described at a high level in Section 2.3.1 and in more detail in Section 4.8.
- The focus of this contribution is on the next variant of expressivity. While the first variant is about capturing a solving circuit, the second variant is about the ability to provide a solution itself. For this purpose, the unitary matrices generated by the parameterized quantum circuit are immediately applied to a fixed initial state, which results in another state, i.e., an element of the unit sphere in some Euclidean space $\mathbb{R}^n$. Any solution of the problem to be solved by the variational quantum algorithm is an element of this unit sphere. Thus, if the set of elements of the unit sphere covered by this process is a significant subset of the unit sphere, the likelihood is high that a solution is reached. The amount of coverage of the unit sphere is assessed by means of the dimension of the submanifold generated by this process [7]. This is what we call the "state approach"; Section 2.3.2 introduces this approach and provides corresponding definitions, while Section 5 gives more details.

The issues with the "state approach" dealt with in this contribution are twofold: Readers of the corresponding existing quantum computing literature require a lot of background from differential topology. We compile and explain this background in detail and in a vivid as well as pedagogical manner; precise references to the matching mathematical literature are given. Furthermore, the corresponding quantum computing literature makes a bunch of statements and claims that are often not proven, or proofs are only indicated. This makes reading and comprehending the literature difficult. We give detailed proofs for claims made in the context of the "state approach"; especially, we prove that the state approach produces in many situations locally immersed submanifolds and even embedded submanifolds (Section 5.2) such that speaking about "dimension" is meaningful at all. We also give a counter-example that these immersions and embeddings can, in general, not be extended to global immersions and embeddings (Section 5.3). Finally, we provide a "recipe" for determining how the local submanifolds can be enlarged (Section 5.4).

Structure of the Article

The article is structured as follows: Section 2 motivates the need for a measure of expressivity (Section 2.2) based on the structure and use of variational quantum algorithms (Section 2.1). The two major definitions of expressivity are given and discussed (Section 2.3); other approaches are briefly reviewed (Section 2.3.3). Section 3 proves that parameterized quantum circuits are differentiable maps, reminds the format of such circuits in many practical cases, and comments on the structure of the parameter space in practical situations. These sections, as well as the next section, are a tutorial of the corresponding subjects.

Section 4 is the main pedagogical part: After proving a lemma on the existence of open chains in connected spaces (which is needed in Section 5.3), the concept of differentiable manifolds with boundaries is introduced, as well as some of their properties that are needed in the rest of this paper (Section 4.2). Section 4.3 discusses singularities that require special care in making claims or assumptions in our context. Next, differentiable maps and

their differentials are discussed (Sections 4.4 and 4.5). Properties of maps with constant rank are reviewed, and immersions as well as embeddings are discussed (Section 4.6). Submanifolds (both immersed as well as embedded) are the subject of Section 4.7. In Section 4.8, the notion of the volume of manifolds is motivated by describing the concept of both linear approximations of manifolds themselves (Section 4.8.1) and "classical" volumes of parallelepipeds (Section 4.8.2). Riemannian manifolds and their volume forms are introduced in Section 4.8.3. Curvature is revealed as the origin of non-uniform distribution of point sets on a manifold and the use of the Haar measure as a solution allowing uniformly random distributions (Section 4.8.6). This allows us in Section 4.8.7 to introduce a precise definition of expressivity in the unitary approach.

Section 5 provides proofs supporting claims from the literature about the state approach. This section contains mostly original contributions. It is proven in Section 5.1 that parameterized quantum circuits induce local immersions. This implies that slices of the parameter space are locally mapped to immersed and even embedded submanifolds the dimensions of which are determined by the local rank of the map induced by the parameterized quantum circuit (Section 5.2). A usual technique to expand local properties to global ones is sketched in Section 5.3, and it is shown that this technique fails in our case, i.e., that local embeddings cannot be extended onto connected components. But a high-level proceeding is roughly sketched in Section 5.4 (together with an example) on how to determine "large" areas of the parameter space that are mapped to an embedded submanifold by the parameterized quantum circuit. The contribution ends with a conclusion in Section 6.

2. The Notion of Expressivity

This chapter discusses the basic principles of variational quantum algorithms (Section 2.1), their origin (Section 2.2.1), and the need for a metric to assess the "success probability" of such algorithms (Section 2.2.2). Two approaches for such a metric—i.e., *expressivity*—are introduced in Section 2.3, and other approaches are briefly reviewed.

2.1. Variational Quantum Algorithms

A *Variational Quantum Algorithm* (VQA) is a pair consisting of a parameterized quantum circuit (also known as. ansatz) and an optimizer [2,8]. The ansatz prepares a state on a quantum computer that is measured [9]. The measurement result is passed to the optimizer that uses a problem-specific cost function to determine, based on the measured values, new parameters for the ansatz. The goal is to find parameters that produce a quantum state whose measurement result optimizes the cost function.

Figure 1 depicts the structure and main ingredients of a variational quantum algorithm. (1) The *parameterized quantum circuit* (PQC) $\mathcal{C}(p_1, \ldots, p_k)$ is given by a unitary operator $\mathcal{C}$ that depends on real parameters $p_1, \ldots, p_k \in \mathbb{R}$. (2) The measurement of the quantum state produced by the quantum circuit is performed by means of problem-specific observables $\{ O_1, \ldots, O_m \}$. The result of the measurement is, thus, (3) a tuple of values $(\langle O_1 \rangle, \ldots, \langle O_m \rangle)$, which is exactly the value of a cost function F in the parameters $p_1, \ldots, p_k$:

$$F(p_1, \ldots, p_k) = (\langle O_1 \rangle, \ldots, \langle O_m \rangle) \tag{1}$$

Thus, a problem is a suitable candidate for being solved on a quantum computer if the cost function cannot be efficiently computed on a classical computer but can be on a quantum computer. Otherwise, a quantum computer would not be needed at all.

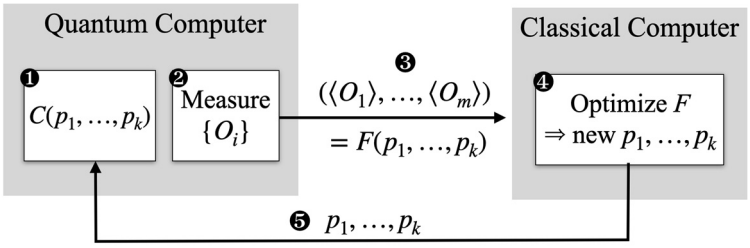


Figure 1. Structure of a VQA.

(4) Next, an optimizer is used: it computes (5) new parameters $p_1, \ldots, p_k$ that move the value of the cost function closer to an optimum. Any kind of optimizer can be used, e.g., gradient-based or derivative-free [10]. When the new parameters have been computed, they are fed into the parameterized quantum circuit $\mathcal{C}(p_1, \ldots, p_k)$, and the loop starts over again (1). The optimizer uses some termination condition, like the number of iterations performed or the amount of improvement achieved by the new parameters. Once the termination criterion is satisfied, the processing stops with the optimal parameters. The quantum circuit is run with the optimal parameters, and measurement of the prepared state in the computational basis represents the solution of the problem at hand.

A Quantum Neural Network (QNN) is also a parameterized quantum circuit [4]. Its cost function depends on training data in addition to other parameters. Once the optimal parameters have been determined (also known as the QNN has been trained), the parameterized quantum circuit is run with the optimal parameters on new data input to perform the task of the QNN (e.g., classification).

Also, VQAs have been applied in several domains: ref. [2] describe applications in multiple areas like finding ground states of molecules, simulating the dynamic evolution of quantum systems themselves, solving linear equation systems (see also [11]), performing classification tasks, or realizing autoencoders. Optimization problems can be addressed, e.g., for energy distribution [12]. Tasks from computational fluid dynamics are addressed [13].

Note that the overall processing of determining the optimal parameters iterates between a quantum part performed on a quantum computer and a classical optimization part performed on a classical computer. Such a kind of processing is called *hybrid quantum-classical* processing, and the corresponding algorithm is called a hybrid quantum-classical algorithm (or just "hybrid" for short). In fact, most quantum algorithms and applications exploiting quantum resources are hybrid [14,15].

Examples of variational quantum algorithms and their use can be found in [16] or [17], for example.

2.2. The Origin of Expressivity

This section provides details about how to assess the suitability of a variational quantum algorithm.

2.2.1. Suitability for NISQ

Variational quantum algorithms have a lot of properties that are of principal interest, e.g., they are universal under certain conditions [18]. Also, some problems are inherently solved by variational quantum algorithms like quantum neural networks [5]. But the main operational aim behind variational quantum algorithms is to achieve an advantage or a utility of quantum computing in the NISQ era (Noisy Intermediate Scale Quantum). This era is roughly defined as the years during which quantum computers are still noisy and have only up to a few hundred qubits [3]. "Noisy" means that the states of these qubits are stable only for a short period of time ("decoherence"), and the gates show small errors

("lack of fidelity"). In this era, implementing quantum applications has a lot of issues to consider [14].

One key consequence of being noisy is that an algorithm must use a QPU only for a short period of time, i.e., less than the decoherence time T_D. A variational quantum algorithm, thus, strives towards an ansatz $\mathcal{C}(p_1, \ldots, p_k)$ and measurement M via observables $\{O_m\}$ such that the time $T_\mathcal{C}$ that the ansatz is executed plus the time T_M that the measurement requires is significantly smaller than the decoherence time:

$$T_\mathcal{C} + T_M \ll T_D \tag{2}$$

2.2.2. Finding Solutions

The other fundamental aim is that the ansatz is capable, in principle, of finding a solution to the problem at hand. This capability can be assessed based on the following:

(i) The plethora of unitaries that an ansatz can produce;
(ii) The plethora of states that an ansatz can produce.

The first approach (that we call the "Unitary Approach" in what follows) considers the parameterized quantum circuit $\mathcal{C}(p_1, \ldots, p_k)$ as a map Ω that maps the parameter space $\mathbb{P}$ of all possible parameters $p_1, \ldots, p_k$ to the unitary group $U(n)$ based on the fact that each $p_1, \ldots, p_k \in \mathbb{P}$ results in a unitary map $\mathcal{C}(p_1, \ldots, p_k) \in U(n)$ (see [19]).

The second approach (called "State Approach" in the following) fixes an initial state $|s\rangle$ and uses for each $p_1, \ldots, p_k \in \mathbb{P}$ the unitary map $\mathcal{C}(p_1, \ldots, p_k) \in U(n)$ to result in an element $\mathcal{C}(p_1, \ldots, p_k)\,|s\rangle \in \mathbb{S}^{n-1}$, which is a map Λ from the parameter space $\mathbb{P}$ to the unit sphere $\mathbb{S}^{n-1}$ (see [7]). Figure 2 depicts both approaches.

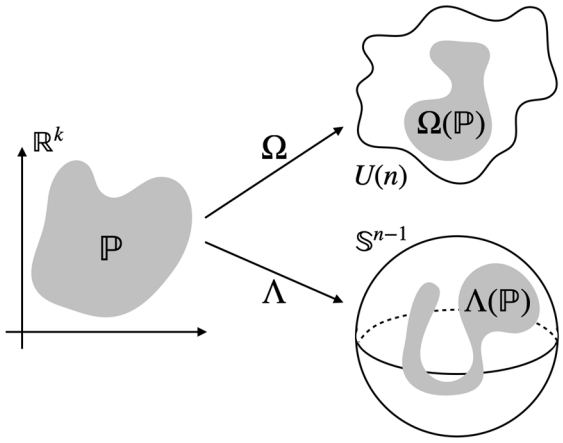


Figure 2. Two Approaches to Assess Expressivity.

Then, a measure of the plethora of unitaries $\Omega(\mathbb{P}) \subseteq U(n)$ or a measure of the plethora of states $\Lambda(\mathbb{P}) \subseteq \mathbb{S}^{n-1}$, respectively, is referred to as the *expressivity* of the ansatz (or of the parameterized quantum circuit, respectively). Thus, in addition to finding an ansatz that satisfies Equation (2), an ansatz should show high expressivity. Finding such an ansatz is considered to be quite difficult (see [9], for example).

Finally, it should be noted that the notion of "state" used here is an abstract one. Such an abstract state is not always related to a state that can be directly realized physically. A reader interested in this subject is referred to [20].

2.3. Defining Expressivity

Expressivity (sometimes also called "expressibility") of an ansatz is measured, for example, by its ability to generate states that represent the Hilbert space well (see [6]), i.e., to uniformly explore the entire Hilbert space (see [2]). In this section, we define both

approaches sketched in the section before more precisely. Note that other definitions exist, but they are often tailored toward specific questions (e.g., see [21]); Section 2.3.3 provides a corresponding brief overview.

2.3.1. The "Unitary Approach"

Let $\mathbb{P} \subseteq \mathbb{R}^k$ be a parameter space and let $\mathcal{C}$ be the ansatz that depends on the parameters $p_1, \ldots, p_k \in \mathbb{P}$. The map $\Omega_\mathcal{C}$ is defined by $\Omega_\mathcal{C} : \mathbb{P} \to U(n)$ with $p_1, \ldots, p_k \mapsto \mathcal{C}(p_1, \ldots, p_k)$. Let X be a problem to be solved by a quantum circuit, and let $\mathcal{S}(X) \subseteq U(n)$ be the set of unitary operators that solve the problem X, i.e., that produce a state that approximately optimizes the corresponding cost function.

If $\Omega_\mathcal{C}(\mathbb{P})$ contains an element of $\mathcal{S}(X)$ the ansatz $\mathcal{C}$ is called *complete* for the problem X. Part (a) of Figure 3 shows the solutions $\mathcal{S}(A)$ and $\mathcal{S}(B)$ of two problems A and B; the ansatz $\mathcal{C}$ is complete for problem B but not complete for problem A.

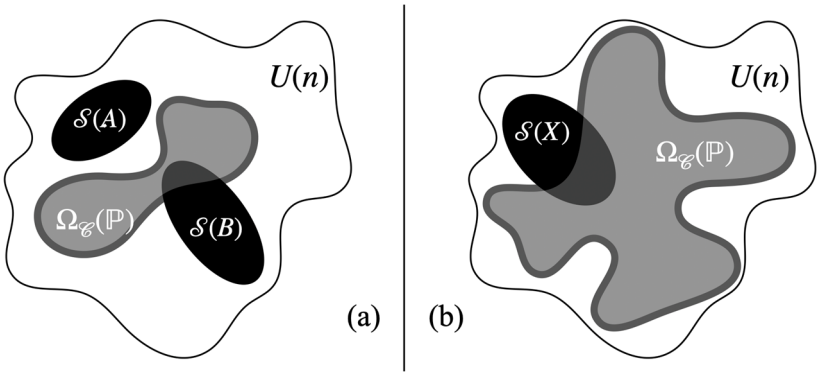


Figure 3. (**a**) Completeness and (**b**) Expressiveness of an Ansatz: The Unitary Approach.

But the solutions $\mathcal{S}(X)$ of a problem X are not known in advance. Thus, the completeness of an ansatz cannot be assessed a priori. Consequently, the more of the unitary group $U(n)$ an ansatz covers (intuitively, "the larger" $\Omega_\mathcal{C}(\mathbb{P}) \subseteq U(n)$ is) the higher the likelihood that it will intersect the set of solutions $\mathcal{S}(X)$.

More precisely, based on the fact that the unitary group $U(n)$ is a Riemannian manifold (see Note 10), a means to measure volumes is available on $U(n)$. Roughly, for an ansatz $\mathcal{C}$, the number $vol(\Omega_\mathcal{C}(\mathbb{P}))$ (i.e., the volume of $\Omega_\mathcal{C}(\mathbb{P})$) is called the *expressivity* of $\mathcal{C}$. If $\mathcal{C}$ and $\widehat{\mathcal{C}}$ are two ansatzes, ansatz $\mathcal{C}$ is called more expressive than ansatz $\widehat{\mathcal{C}}$ if $vol(\Omega_\mathcal{C}(\mathbb{P})) \supset vol(\Omega_\mathcal{C}(\mathbb{P}))$.

Note that this definition is not the definition used in the literature, but it has the purpose of being descriptive: the exact definition of expressivity of an ansatz $\mathcal{C}$ based on the "unitary approach" compares its distribution of states from sampling with the collection of Haar-random states (see [4,6]). In Section 4.8, we provide more background on this.

2.3.2. The "State Approach"

This approach is the one we focus on mostly; thus, we define it precisely.

Definition 1. *Let $\mathcal{C}$ be an ansatz depending on parameters $p_1, \ldots, p_k$. The set of all parameters $\mathbb{P} \subseteq \mathbb{R}^k$ allowed by $\mathcal{C}$ is called the parameter space of $\mathcal{C}$. Let $|\iota\rangle \in \mathbb{S}^{n-1}$ be a fixed chosen state, the so-called initial state. Then, the map $\Lambda_\mathcal{C}$ defined by $\Lambda_\mathcal{C} : \mathbb{P} \to \mathbb{S}^{n-1}$ with $p_1, \ldots, p_k \mapsto \mathcal{C}(p_1, \ldots, p_k)\,|\iota\rangle$ is called the state map of the parameterized circuit $\mathcal{C}$.*

$\Lambda_\mathcal{C}(\mathbb{P})$ is the set of all states reachable by means of the parameterized circuit $\mathcal{C}$. As proven in Section 5.1, $\Lambda_\mathcal{C}(\mathbb{P})$ is under practical conditions locally an immersed submanifold (Lemma 16) of $\mathbb{S}^{n-1}$, and locally often even an embedded submanifold of $\mathbb{S}^{n-1}$ (Lemma 17).

The reader is referred to Sections 4.2 and 4.7 (or to textbooks like [22] or [23], respectively) for the definition of the terms "manifold" and "submanifold" (often, "manifolds" are assumed to be smooth, i.e., of class C^∞, in what follows—which is without loss of generality according to Theorem 1).

Figure 4 depicts the images of the parameter space $\mathbb{P} \subseteq \mathbb{R}^k$ in the unit sphere $\mathbb{S}^{n-1}$ produced by two different ansatzes $\mathcal{C}$ and $\widetilde{\mathcal{C}}$. In case a circuit produces a submanifold of $\mathbb{S}^{n-1}$ (immersed or embedded—see Section 5), this submanifold is called a *circuit manifold*, i.e., the circuit manifold is the set of states reachable by means of the parameterized circuit. In the figure, the circuit manifold $\Lambda_{\mathcal{C}}(\mathbb{P})$ has dimension 1, and the circuit manifold $\Lambda_{\widetilde{\mathcal{C}}}(\mathbb{P})$ is of dimension 2. Intuitively, the dimension of a circuit manifold is a measure of the expressivity of a parameterized circuit [7]: the intuition is that a higher-dimensional circuit manifold covers more of the enclosing manifold and, thus, has a higher chance of meeting the solution of a problem.

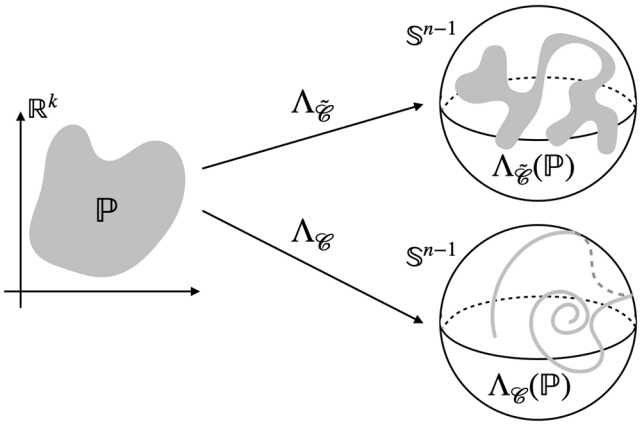


Figure 4. Expressiveness of an Ansatz: The State Approach.

But this intuition is misleading: for example (see [23], example 4.20), there is an immersed submanifold L of the torus $\mathbb{T}^2$ which is dense in $\mathbb{T}^2$, but $dim L = 1$. i.e., L is arbitrarily close to any point $p \notin L$. Thus, a circuit manifold of dimension 1 may approximate a solution with arbitrary precision (see Figure 5b for a rough indication), while a "small" higher-dimensional circuit manifold may stay "far away" from the solution (see Figure 5a). This means that, in general, dimensional expressivity is not a reliable indicator for the success of an ansatz: an ansatz with a circuit manifold with maximum dimensional expressivity might not solve the underlying problem, while an ansatz with a lower dimensional expressivity might be successful.

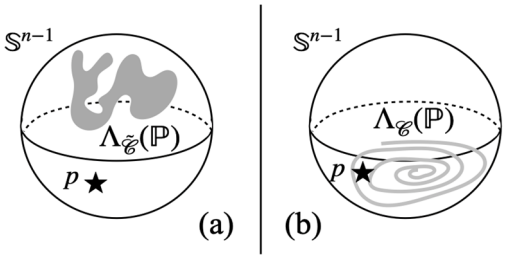


Figure 5. Misleading Intuition of Dimensional Expressivity.

However, the dimensional expressivity allows us to identify parameters of a parameterized circuit that are superfluous, i.e., that can be removed without affecting the capability of an ansatz to determine a solution (see [7] for the details). A reduced number of parameters has the advantage of a speedup of the overall processing, e.g., by less optimization work required.

2.3.3. Other Approaches and Relations

While the two approaches sketched before are generally applicable, other approaches to assess expressivity have been proposed, which are specific to certain domains. Also, the relation between expressivity and other aspects of a circuit has been investigated.

For example, ref. [24] introduces a measure of expressivity called the covering number, which is derived from statistical learning theory. This number is the minimum number of balls of a given radius that covers the hypothesis space of a parameterized quantum circuit. Here, the hypothesis space is the subset of unitaries made available by using all parameters. An estimation of the covering number is given that shows that this measure of expressivity depends on the number of qubits used by the circuit, its gates, its observables, as well as the error rate of the quantum computer used.

While the estimation of [24] provides an upper bound of the covering number of a parameterized quantum circuit, ref. [25] derives a lower bound of the covering number. This can be used to determine the appropriateness of a given circuit.

Another measure of expressivity is the effective rank introduced in [26]. The effective rank, i.e., the rank of the quantum Fisher information matrix [27] of a circuit, represents the number of independent parameters of a parameterized quantum circuit. It can be used to determine "superfluous" parameters that can be eliminated, thus, making optimization in the corresponding variational quantum algorithm more effective.

Expressivity can also be assessed by the ability of a parameterized quantum circuit to approximate certain classes of functions. For example, ref. [28] investigates quantum neural networks that approximate multivariate polynomials and Hölder smooth functions. Their notion of expressivity is how well a parameterized quantum circuit can approximate such functions. They explicitly construct such circuits and give approximation boundaries in terms of depth and width of the constructed circuit.

Ref. [29] focusses on parameterized quantum circuits that can be represented as a Fourier series. This is the case when the circuit not only depends on parameters that are iteratively modified but also on input data that is reflected in the circuit by matrix exponentials of Hamiltonian operators. The spectrum of the Fourier series then depends on the eigenvalues of these Hamiltonians, and the Fourier coefficients depend on the parameterized unitaries of the circuit. Ref. [29] uses the unitary approach sketched in Section 2.3.1 to measure expressivity but shows that such parameterized quantum circuits suffer from exponentially vanishing Fourier coefficients for growing numbers of qubits. This limits the expressivity of the corresponding circuits.

Using the unitary approach, ref. [30] studies the expressivity of a parameterized quantum circuit and its entangling power in relation to specific classes of gates used within the circuit. The main finding is that the use of so-called quantum switches in a circuit result in better expressivity when the depth of the circuit increases. Also, the corresponding circuits result in higher entanglement power when their depth increases.

The impact of the types of gates used in a parameterized quantum circuit on its expressivity (according to the unitary approach) is also studied in [31]. It turns out that CNOT gates have a decreasing effect on expressivity, while rotational RX, RY, and RZ gates strengthen it: RX has the strongest effect, followed by RY, followed by RZ.

Ref. [32] investigates the impact of the graph structure of a quantum computer—i.e., the graph consisting of the qubits of the quantum computer as nodes and the physical connections between the qubits as edges—on the expressivity of circuits executed on the machine. Based on two different ansatzes of parameterized quantum circuits, their finding is that a graph with a ring structure achieves the highest expressivity (measured via the unitary approach), followed by a linear graph, an any-to-any graph, and a star-shaped graph.

Variational quantum algorithms are subject to so-called barren plateaus: the graph of the cost function C may become flat such that the optimizer can no longer efficiently determine better parameters. This is a significant problem limiting the practical use of variational quantum algorithms. It turns out that high expressivity is one possible source of barren plateaus. Ref. [33] presents a Lie algebraic theory that can be used to assess the advent of barren plateaus. This allows us to determine whether the expressivity must be adapted.

3. Parameterized Quantum Circuits as Differentiable Maps

Expressivity depends on properties of parameterized circuits and their images in $U(n)$ or $\mathbb{S}^{n-1}$. The proofs of these properties are mostly sketched only in the literature. In what follows, these proofs are given in detail.

3.1. Proof of Differentiability

The most fundamental observation is that in many practical situations, a parameterized quantum circuit is a differentiable map (even *smooth*, i.e., of class C^∞). This can be seen by noting that each unitary can be represented by a set of 1-qubit operations and CNOTs, and by proving that these ingredient operations are smooth.

To begin with, we need the following known fact from linear algebra:

Note 1. *For each unitary matrix A there exists a Hermitian matrix H such that $A = e^{iH}$.*

Proof. Each unitary is diagonalizable, i.e., there exists a $T \in U(n)$ such that $A = T diag\left(e^{i\varphi_1}, \cdots, e^{i\varphi_n}\right) T^*$ with real numbers $\varphi_1, \ldots, \varphi_n \in \mathbb{R}$ (see [34] Theorem 18.13). Define the real diagonal matrix $\Phi := diag(\varphi_1, \ldots, \varphi_n)$ and set $H := T\Phi T^*$. Then, H is Hermitian because $H^* = (T\Phi T^*)^* = (T^*)^* \Phi^* T^* \overset{(a)}{=} T\Phi T^* = H$, where (a) is valid because $\Phi^* = \Phi$ for a real diagonal matrix Φ. Then it is

$$\begin{aligned} e^{iH} &= e^{iT\Phi T^*} \\ &\overset{(1)}{=} e^{iT\Phi T^{-1}} \\ &\overset{(2)}{=} T e^{i\Phi} T^{-1} \\ &= T e^{diag(i\varphi_1, \ldots, i\varphi_n)} T^{-1} \\ &\overset{(3)}{=} T diag\left(e^{i\varphi_1}, \ldots, e^{i\varphi_n}\right) T^{-1} \\ &= A \end{aligned}$$

Hereby, (1) is valid because T is unitary, i.e., $I = TT^*$; thus, $T^{-1} = T^*$. (2) is valid because $e^{SXS^{-1}} = Se^X S^{-1}$ for each $S \in GL(n)$, which is seen as follows:

$$\begin{aligned} e^{SXS^{-1}} &= \sum_{k=1}^{\infty} \frac{\left(SXS^{-1}\right)^k}{k!} \\ &= \sum_{k=1}^{\infty} S \frac{X^k}{k!} S^{-1} \\ &= S \left(\sum_{k=1}^{\infty} \frac{X^k}{k!} \right) S^{-1} \\ &= S e^X S^{-1} \end{aligned}$$

Finally, (3) is valid because

$$\begin{aligned} e^{diag(x_1,\ldots,x_n)} &= \sum_{k=1}^{\infty} \frac{diag\left(x_1,\ldots,x_n\right)^k}{k!} \\ &= \sum_{k=1}^{\infty} \frac{diag\left(x_1^k,\ldots,x_n^k\right)}{k!} \\ &= diag(e^{x_1},\ldots,e^{x_n}) \end{aligned}$$

Thus, each unitary matrix is the matrix exponential of a Hermitian matrix. □

Note 2. *Each parameterized 1-qubit operation* $U : \mathbb{R} \supseteq I \to U(2)$ *is smooth.*

Proof. According to Note 1, each unitary matrix U is a matrix exponential of a Hermitian matrix H: $U = e^{iH}$. Especially, for $I \subseteq \mathbb{R}$ and a parameterized 1-qubit operation $U : I \to U(2)$, $p \mapsto U(p)$, there exists a Hermitian matrix H such that $U(p) = e^{ipH}$. Thus, $U(p)$ is a smooth operation. □

If the 1-qubit operation U modifies the i-th qubit of an n qubit quantum register (i.e., all other qubits $j \neq i$ are left unchanged) it can be represented as the following tensor product:

$$T(p) = I \otimes \cdots \otimes I \otimes \overbrace{U(p)}^{i} \otimes I \cdots \otimes I \tag{3}$$

This is a smooth operation because the coefficients of the corresponding matrices are either constant (i.e., 0 or 1 in case the matrix I) or are the coefficients of the smooth 1-qubit operation $U(p)$. Finally, building the tensor product of matrices is a differentiable operation itself.

The same argument is used to prove the following:

Note 3. *The controlled NOT operator with control qubit r and target qubit s* $CX^{[r,s]} : \mathbb{H}^n \to \mathbb{H}^n$ *on an n qubit quantum register is smooth.*

Proof. The matrix $CX^{[r,s]}$ is a permutation matrix, i.e., its coefficients are constant $(CX^{[r,s]})_{ij} \in \{0,1\}$. Thus, the matrix is a smooth operator. □

A parameterized quantum circuit $\mathcal{C}(p_1,\ldots,p_k)$ has a parameter space $\mathbb{P} \subseteq \mathbb{R}^k$ (we will discuss properties of the topological structure of the parameter space in Section 3.3), and for each parameter tuple $(p_1,\ldots,p_k) \in \mathbb{P}$ the resulting operator is a unitary: $\mathcal{C}(p_1,\ldots,p_k) \in U(n)$. Thus, a parameterized quantum circuit is a map

$$\begin{array}{ll} \mathcal{C} : \mathbb{P} & \to U(n) \\ (p_1,\ldots,p_k) & \mapsto \mathcal{C}(p_1,\ldots,p_k) \end{array} \tag{4}$$

As a unitary, $\mathcal{C}(p_1,\ldots,p_k)$ is composed of (parameterized) 1-qubit operators (in the representation of Equation (3)) and CNOT operators $CX^{[r,s]}$ (see [1], Section 4.5.2). This composition is performed by multiplying the corresponding matrices. But matrix multiplication is smooth: if $\left(a_{ij}\right)$, $\left(b_{st}\right)$ are two matrices, then their product is a matrix, the coefficients of which are a sum of the products of the coefficients of the factor matrices, i.e., $\left(a_{ij}\right) \cdot \left(b_{st}\right) = \left(\sum_j a_{ij}b_{jt}\right)_{i,t}$, thus multiplying matrices are smooth. If the factor matrices of the matrix product are smooth (i.e., their coefficients $\left(a_{ij}\right)$, $\left(b_{st}\right)$ are smooth functions), the result is smooth again. Because the factors of the matrix product resulting in $\mathcal{C}(p_1,\ldots,p_k)$

are smooth parameterized 1-qubit operators and smooth CNOTs (see Note 2 and Note 3), $\mathcal{C}(p_1, \ldots, p_k)$ is smooth.

Let $(\mathcal{U}_{ij}(p_1, \ldots, p_k))_{i,j}$ be the matrix representation of $\mathcal{C}(p_1, \ldots, p_k)$. We choose an arbitrary but fixed initial state $|\iota\rangle = |\iota_1, \ldots, \iota_n\rangle \in \mathbb{H}^n$. Then the product $\mathcal{U}(p_1, \ldots, p_k)\,|\iota\rangle = (\sum_j \mathcal{U}_{ij}(p_1, \ldots, p_k) \cdot \iota_j)_i$ is smooth, i.e., the components of the resulting vector are smooth. This finally proves:

Lemma 1. *Let $\mathcal{C} : \mathbb{P} \to U(n)$ be a parameterized quantum circuit with matrix representation $(\mathcal{U}_{ij}(p_1, \ldots, p_k))$ and let $|\iota\rangle \in \mathbb{S}^{n-1}$ be a fixed initial state. Then, both maps $\Omega_{\mathcal{C}} : \mathbb{P} \to U(n)$ with $(p_1, \ldots, p_k) \mapsto \mathcal{C}(p_1, \ldots, p_k)$ as well as $\Lambda : \mathbb{P} \to \mathbb{S}^{n-1}$ with $(p_1, \ldots, p_k) \mapsto \mathcal{U}(p_1, \ldots, p_k)\,|\iota\rangle$ are smooth.*

Figure 6 summarizes the ingredients of a parameterized quantum circuit: The circuit has an associated parameter space $\mathbb{P}$. The parameterized quantum circuit $\mathcal{C}$ itself assigns to each parameter tuple $(p_1, \ldots, p_k) \in \mathbb{P}$ a unitary operator $\mathcal{C}(p_1, \ldots, p_k) \in U(n)$ (we do not distinguish between the circuit $\mathcal{C}$ and its matrix representation $\mathcal{U}$). This unitary operator transforms an initial state $|\iota\rangle \in \mathbb{S}^{n-1}$ into another state $\mathcal{C}(p_1, \ldots, p_k)\,|\iota\rangle \in \mathbb{S}^{n-1}$. Thus, a parameterized quantum circuit $\mathcal{C}$ induces a smooth map $\Lambda : \mathbb{P} \to \mathbb{S}^{n-1}$ that maps its parameter space $\mathbb{P}$ into the unit sphere $\mathbb{S}^{n-1}$.

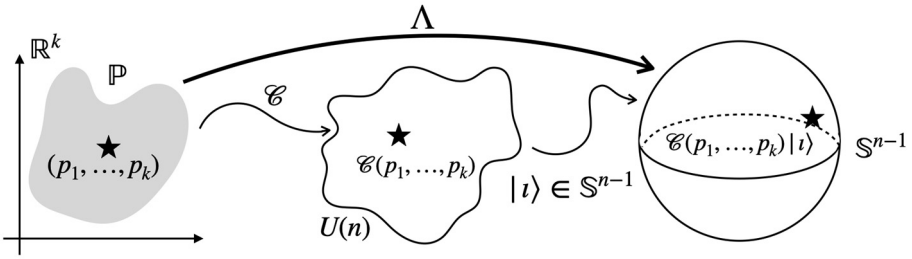


Figure 6. Parameterized Quantum Circuit as a Smooth Map.

3.2. Differentiability in Practice

In addition, an argument from practice can be used to see that a typical parameterized quantum circuit of a variational quantum algorithm is smooth. This is because often (see [2,19]), a typical circuit has the format

$$\begin{aligned} \mathcal{C} : \mathbb{P} &\to U(n) \\ (p_1, \ldots, p_k) &\mapsto \prod_{j=1}^{k} e^{-ip_j H_j} W_j \end{aligned} \tag{5}$$

with a set of fixed unitary operators $\{W_j\}$ and Hermitian operators $\{H_j\}$, i.e., $e^{-ip_j H_j}$ is a rotation of angle p_j generated by H_j. The latter rotations are smooth, their product with a constant matrix is smooth, and, thus, the overall product is smooth. The parameter space $\mathbb{P}$ in this case is a Cartesian product of connected intervals in $\mathbb{R}$ which are the domains of the corresponding angles p_j.

3.3. Differentiability on Parameter Spaces

By now we have not commented on the structure of a parameter space $\mathbb{P}$ (except for the special case of Section 3.2). Typically, a parameter space is not an arbitrary set in $\mathbb{R}^k$ but it is assumed to have suitable properties. Especially, $\mathbb{P}$ must support the notion of maps that are differentiable on $\mathbb{P}$ as their domain (see Section 4.4). For example, for arbitrary sets $\mathbb{P} \subseteq \mathbb{R}^k$ a map $f : \mathbb{P} \to \mathbb{R}^n$ is differentiable, iff a function $F : U \to \mathbb{R}^n$ with $\mathbb{P} \subseteq U \subseteq_{open} \mathbb{R}^k$

exists that is differentiable in the ordinary sense and that fulfills $F|_{\mathbb{P}} = f$, i.e., F restricted to $\mathbb{P}$ is identical to f (in this case, F is called a *differentiable extension* of f)—see Definition 5.

As another example, if $\mathbb{P}$ is a differentiable manifold (with or without boundary), differentiability is defined in every point $x \in \mathbb{P}$ by means of an appropriate chart around x (see Section 4.4 and [23]).

In circuits like the ones in Equation (5), $\mathbb{P}$ is a Cartesian product of open intervals. This can slightly be generalized as follows (see Section 4.2): $\mathbb{P} = \mathbb{P}_1 \times \ldots \times \mathbb{P}_k \subseteq \mathbb{R}^k$ with $\mathbb{P}_i \subseteq_{Mf} \mathbb{R}$ being a connected manifold (with or without boundary) of dimension one, $dim\mathbb{P}_i = 1$. Such a manifold $\mathbb{P}_i$ is an open, closed, or semi-closed interval. Thus, $\mathbb{P}$ is a hypercube where some of its faces may belong to the hypercube; note that not all of the intervals $\mathbb{P}_i$ may include boundary points to avoid singularities—see Section 4.3. Thus, in general, $\mathbb{P}$ is a manifold with boundary (Section 4.2).

4. Facts from (Differential) Topology

In this Section, we prove that two points in a connected topological space are connected by an open chain (Lemma 2). The definition of manifolds with boundaries is given, as well as the differentiability of maps. The concept of singularities is presented (Section 3.3). Differentiability of maps, their differentials, and properties of maps of constant rank are summarized. Also, we remind the rank theorem, which implies that immersions are locally injective, and we remind a condition under which injective immersions are embeddings. Submanifolds are introduced, and their properties are discussed. Section 4.8 introduces the volume form on manifolds as well as the Haar measure; this allows us to present in Section 4.8.7 more details about the "unitary approach" (as sketched informally in Section 2.3.1).

4.1. A Property of Connected Spaces

In this Section, we remind the notion of connectedness of a topological space (refer to the textbook [35] for details). We prove a property of connected spaces (Lemma 2) that we need later and that we did not find in the English literature (but in the German textbook [36]).

Definition 2. *A topological space X is called connected if for two non-empty open sets $O_1, O_2 \subseteq_{open} X$, $O_1 \neq \varnothing$, $O_2 \neq \varnothing$ with $X = O_1 \cup O_2$ it follows that $O_1 \cap O_2 \neq \varnothing$.*

In other words, a topological space is connected if it cannot be split into two disjoint, non-empty open sets. The following is well-known (e.g., [35], Proposition 9.8):

Note 4. *Let X be a connected topological space and let $Z \subseteq X$ be both, open and closed. Then $Z = \varnothing$ or $Z = X$.*

Definition 3. *A family of open sets $\mathcal{U} = \{U_i\}_{i \in I}$, $U_i \subseteq_{open} X$ with $\bigcup_{i \in I} U_i = X$ is called an open cover of X.*

The proof of the following lemma is from [36].

Lemma 2. *Let X be a topological space. X is connected $\Leftrightarrow$ For each open cover $\mathcal{U}$ of X and two points $a, b \in X$ there are $\{U_1, \ldots, U_n\} \subseteq \mathcal{U}$ with:*

(i) $a \in U_1$, $a \notin U_i$ *for* $i \neq 1$
(ii) $b \in U_n$, $b \notin U_i$ *for* $i \neq n$
(iii) $U_i \cap U_j \neq \varnothing \Leftrightarrow |i - j| \leq 1$

An open cover $\mathcal{U}$ with the properties (i), (ii), and (iii) for each pair of points $a, b \in X$ is called an *open chain* connecting the points $a, b \in X$ (see Figure 7).

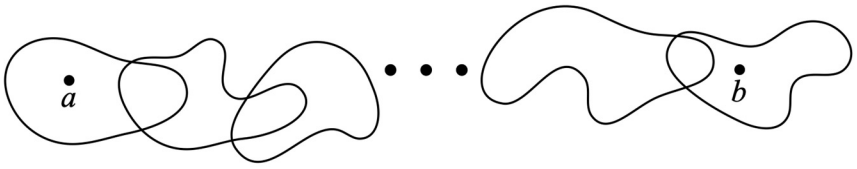


Figure 7. Chain of Open Sets Connecting Two Points.

Proof. "⇐" Let X be disconnected. We must show that there is an open cover of X that does not satisfy conditions (i), (ii), and (iii). Because X is not connected, we find open sets $O_1, O_2 \subseteq_{open} X$ with $O_1 \neq \varnothing, O_2 \neq \varnothing, O_1 \cap O_2 = \varnothing$, and $O_1 \cup O_2 = X$. Choose $a \in O_1$ and $b \in O_2$. The open cover $\mathcal{U} = \{O_1, O_2\}$ satisfies (i) and (ii) but not (iii).

"⇒" Let $\mathcal{U}$ be an open cover of X. Two points $a, b \in X$ are called $\mathcal{U}$-connected (in symbols $a \sim b$) if there exists $\{U_1, \ldots, U_n\} \subseteq \mathcal{U}$ with the properties (i), (ii), and (iii).

Then, "$\sim$" is an equivalence relation: reflexivity and symmetry are obvious. Transitivity is seen as follows: for $a, b, c \in X$ with $a \sim b$ and $b \sim c$ choose a chain of open sets $\{U_1, \ldots, U_n\} \subseteq \mathcal{U}$ connecting a, b and choose a chain of open sets $\{V_1, \ldots, V_m\} \subseteq \mathcal{U}$ connecting b, c.

Set $r := min\{i \in \{1,\ldots,n\} \mid \exists 1 \leq j \leq m : U_i \cap V_j \neq \varnothing\}$ and $s := max\{j \in \{1,\ldots,m\} \mid U_r \cap V_j \neq \varnothing\}$. Then, $\{U_1, \ldots, U_r, V_s, \ldots, V_m\} \subseteq \mathcal{U}$ is a chain of open sets with properties (i), (ii), (iii) connecting $a,\ c$, i.e., $a \sim c$.

Let $[x] \subseteq X$ be a $\sim$-equivalence class. If we show that $[x] = X$ then we are done. For $y \in [x]$ and a chain of open sets $\{U_1, \ldots, U_n\} \subseteq \mathcal{U}$ connecting x, y it is $U_n \subseteq_{open} X$ by definition, and it is $U_n \subseteq [x]$ because each $z \in U_n$ is $\mathcal{U}$-connected with x. Thus, $[x] \subseteq_{open} X$. This implies that $X \setminus [x] = \bigcup_{y \notin [x]} [y] \subseteq_{open} X$, thus $[x] \subseteq_{closed} X$.

Thus, $[x] \subseteq X$ is open and closed in X, but $[x] \neq \varnothing$. Note 4 before implies $[x] = X$. □

4.2. Differentiable Manifolds

The definition of a manifold with boundaries is given (the seminal textbook [37] provides many details and proofs, and [23] is a more modern treatment of the subject). For this purpose, we need the definition of the n-dimensional real half-space $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n | x_n \geq 0\}$; it is $\partial\mathbb{R}^n_+ = \{x \in \mathbb{R}^n | x_n = 0\} \cong \mathbb{R}^{n-1}$ the *boundary* of $\mathbb{R}^n_+$.

Note that in the context of differentiable manifolds, topological spaces are considered to be Hausdorff spaces and second countable. A *Hausdorff space* requires that any two different points of the space have disjoint neighborhoods; the set of all neighborhoods of a point p is denoted by $\mathfrak{U}_p$. A space is *second countable* if it has a countable basis, i.e., every open set of the space is the union of a subset of the basis. See [38] for the detailed definitions of these terms.

Definition 4. *Let M be a Hausdorff and second countable topological space. If for each point $x \in M$ there exists an open neighborhood $U \subseteq M$ and a homeomorphism $\varphi : U \to \varphi(U) \subseteq_{open} \mathbb{R}^n_+$, then M is called a topological manifold with boundary of dimension n (in symbol: $dim M = n$). The pair (U, φ) is called a chart of M around x. A set of charts $\mathfrak{A} = \{(U_i, \varphi_i) | i \in I\}$ with $\bigcup_{i \in I} U_i = M$ is called an atlas of M.*

A point $x \in M$ with $\varphi(U) \subseteq_{open} \mathbb{R}^n$ is called interior point, and a point $x \in M$ with $\varphi(x) \in \partial\mathbb{R}^n_+$ is called boundary point. The set of boundary points of M is called the boundary of M, denoted by ∂M.

For two intersecting charts $(U_i, \varphi_i), (U_j, \varphi_j) \in \mathfrak{A}$, i.e., *charts with $U_i \cap U_j \neq \varnothing$ the map $\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$ is called the transition function between the charts.*

M is called a differentiable manifold of class C^r (C^r-manifold for short) if all transition functions are differentiable of class C^r ; the corresponding atlas $\mathfrak{A}$ is called C^r-atlas. For $r = \infty$, the manifold (and the atlas) is called smooth.

In Figure 8 p and q are interior points while r is a boundary point.

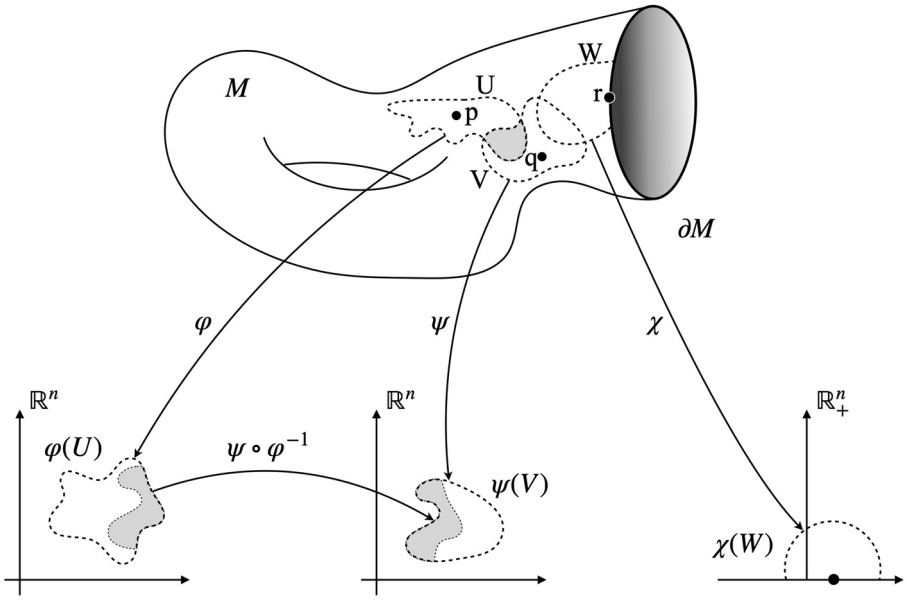


Figure 8. Charts and Transition Functions of a Differentiable Manifold With Boundary.

Textbooks on differential topology (e.g., [23]) or on differential geometry (e.g., [39]) contain many examples of differentiable manifolds as well as corresponding atlases. Standard examples include n-dimensional spheres in the (n + 1)-dimensional Euclidean space $\mathbb{R}^n$, a torus, graphs of differentiable functions (see Lemma 9), which then include curves and surfaces in $\mathbb{R}^3$. Especially, the Euclidean space $\mathbb{R}^n$ is a differentiable manifold (see next).

A C^r-manifold may have several C^r-atlases. For example, both, $\mathfrak{A}_1 = \left\{\left(\mathbb{R}^k, id\right)\right\}$ as well as $\mathfrak{A}_2 = \left\{\left(\mathbb{R}^k_{x_k>-1}, id\right), \left(\mathbb{R}^k_{x_k<1}, id\right)\right\}$ are C^r-atlases (for any r) of $\mathbb{R}^k$, and so is $\mathfrak{A}_1 \cup \mathfrak{A}_2$. In general, when adding a chart (U, φ) to a given C^r-atlas $\mathfrak{A}$ and the resulting atlas $\mathfrak{A} \cup \{(U, \varphi)\}$ is again a C^r-atlas, (U, φ) is said to be *compatible* with $\mathfrak{A}$. Adding all compatible charts to $\mathfrak{A}$ results in the (unique) maximal atlas $\overline{\mathfrak{A}}$ (that contains $\mathfrak{A}$). A maximal C^r-atlas is called a *C^r-differentiable structure* of the corresponding manifold. Any C^r-atlas $\mathfrak{A}$ determines a unique C^r-differentiable structure (see [23], Proposition 1.17). Thus, we can assume that the atlases of a manifold are differentiable structures.

The following is often used, and its proof can be found in [37].

Lemma 3. *Let M be a C^r-manifold with boundary, $dim M = n$. Then, $\partial M \subseteq_{closed} M$ and ∂M is a C^r-manifold without boundary, $dim \partial M = n - 1$.*

Also, it is well-known that any open subset of a manifold is again a manifold:

Note 5. *(a) $\mathbb{R}^k$ is a smooth manifold without boundary.*
(b) For a C^r-manifold M any open subset $S \subseteq_{open} M$ is a C^r-manifold.
(c) The interior $M \setminus \partial M$ of a manifold with boundary is a manifold without boundary.

Proof. (a) $\mathfrak{A} = \left\{\left(\mathbb{R}^k, \mathrm{id}\right)\right\}$ is a smooth atlas that shows that $\mathbb{R}^k$ is a manifold without a boundary.

(b) If (U, φ) is a chart of M around $x \in S$, then $x \in U \cap S$ and $\varphi|_{U \cap S}$ is a homeomorphism. Thus, $(U \cap S, \varphi|_{U \cap S})$ is a chart of S around x. Also, restrictions of transition functions are of the same differentiability class as the original functions.

(c) It is $M \setminus \partial M \subseteq_{open} M$ and, thus, the claim follows from (b). □

There are several ways in which manifolds can be constructed from existing manifolds, e.g., by building their sums, quotients, and products (see [23]). In our context, the product of manifolds is of interest: if $\{M_1, \ldots, M_k\}$ are C^r-manifolds (without boundary) and $dim M_i = n_i$, then their Cartesian product $M_1 \times \ldots \times M_k$ is another C^r-manifold (without boundary). For example,

- The unit sphere $\mathbb{S}^1 \subseteq \mathbb{R}^2$ is a smooth manifold. Thus, $\mathbb{T}^n := \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ (the product of n copies of $\mathbb{S}^1$) is a new smooth manifold called an n-dimensional torus.
- An open interval $]a_i, b_i[\subseteq_{open} \mathbb{R}$ (for $1 \leq i \leq n$) is a smooth manifold. Thus, the open n-dimensional cuboid $Q^n :=]a_1, b_1[\times \cdots \times]a_n, b_n[\subseteq \mathbb{R}^n$ is a smooth manifold.

For manifolds with boundaries, building their products is a bit less straightforward. In general, there is an extensive theory behind this (see [40]). In our context, the following facts suffice:

1. Let $\{M_1, \ldots, M_k\}$ be C^r -manifolds (without boundary), $dim M_i = n_i$, and let N be a C^r-manifold with boundary, $dim N = n$. Then, $M_1 \times \cdots \times M_k \times N$ is a manifold with boundary, $dim(M_1 \times \cdots \times M_k \times N) = n_1 + \cdots + n_k + n$, and $\partial(M_1 \times \cdots \times M_k \times N) = M_1 \times \cdots \times M_k \times \partial N$ (see [23], Proposition 1.45).
2. Thus, with $]a_i, b_i[\subseteq_{open} \mathbb{R}$ for $1 \leq i \leq n$ and $[a, b] \subseteq_{closed} \mathbb{R}$, the product $Q :=]a_1, b_1[\times \cdots \times]a_n, b_n[\times [a, b] \subseteq \mathbb{R}^{n+1}$ is a manifold with boundary $\partial Q =]a_1, b_1[\times \cdots \times]a_n, b_n[\times \{a\} \cup]a_1, b_1[\times \cdots \times]a_n, b_n[\times \{b\}$ (for an example, see Figure 9).

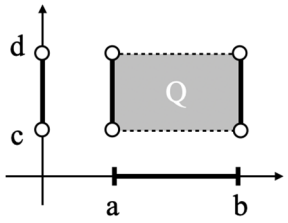


Figure 9. A Product Manifold Q With Boundary.

4.3. Singularities

In Figure 9, the boundary of the manifold Q consists of the left and right edges without their endpoints. This is an implication of the fact that a boundary of a manifold with boundary is a manifold without boundary (see Lemma 3). If the endpoints of the left or right edges were included, they would become the boundary of the manifolds consisting of the left or right edges: contradiction.

This is a general phenomenon: manifolds exclude *singularities*, i.e., non-differentiable structures. A vertex of a rectangle is an example of such a singularity: Part (a) in Figure 10 depicts the rectangle Q and one of its vertices x. Part (b) focuses on a neighborhood of x in Q. Finally, part (c) extracts the boundary of this neighborhood, moves x to the origin, and rotates the boundary by 45°; movements and rotations are smooth maps, i.e., the resulting graph in (c) is diffeomorphic to the part of the boundary in (b). Obviously, the graph in (c) is the graph of the absolute function. Assume that there is a chart (U, φ) of the boundary around x with $\varphi(x) = 0$. Then φ is a diffeomorphism (see Note 6 below); thus, φ^{-1} is differentiable, but φ^{-1} is the absolute function, which is known to be not differentiable

at 0. This contradiction shows that x is a singularity, i.e., vertices must not be part of a rectangular "shape" in order to be a manifold.

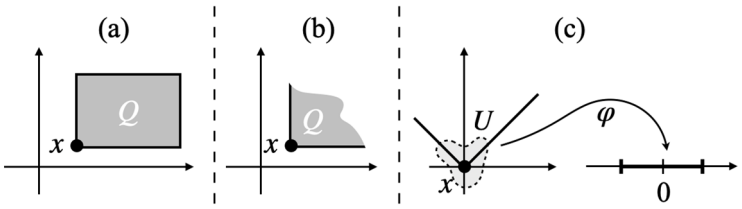


Figure 10. Vertices in Rectangles are Singularities.

Many other kinds of singularities exist. For example (see Figure 11): Part (a) of the figure shows a "cusp", which is a curve that is not differentiable at the point x. An intersection at point y in part (b) is not differentiable. Also, an edge of a cuboid (see part (c) of the figure) consists of nodes, each of which is a singularity.

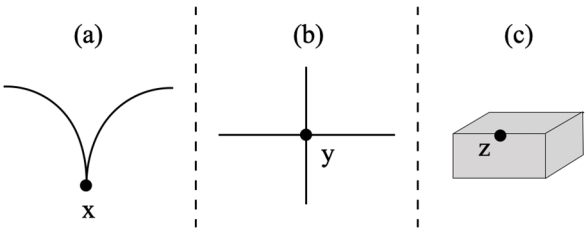


Figure 11. More Singularities.

The intersection point in part (b) is already a topological singularity; i.e., no reference to differentiability is needed to recognize this: Any point different from y is contained in a neighborhood that is homeomorphic to an open interval in $\mathbb{R}$. Thus, if a chart $(U,\ \varphi)$ around y would exist, U could be chosen to be connected. Then, $\varphi(U) \subseteq \mathbb{R}$ is connected. $U \setminus \{y\}$ consists of four connected components. Deleting a single point from a connected subset of the real line results in two connected components, i.e., $\varphi(U) \setminus \{\varphi(y)\} = \varphi(U \setminus \{y\})$ consists of two connected components—a contradiction because $\varphi(U \setminus \{y\})$ consists of four connected components.

Many of the results achieved in differential topology of manifolds are not valid in the presence of singularities. This is why special care must be taken when claiming that geometric objects are manifolds (with or without boundary): it must be proven that they are manifolds to avoid singularities. Singularities are extensively studied (see [41,42], or [43] for example).

4.4. Differentiable Maps

In Figure 8, $\chi(W)$ is not an open set in $\mathbb{R}^n$; thus, the ordinary definition of differentiability (that is typically defined for open sets) does not apply. Consequently, the notion of differentiability is extended to functions with a domain that is an arbitrary subset of $\mathbb{R}^n$.

Definition 5. *Let $A \subseteq \mathbb{R}^n$ be an arbitrary set and let $f : A \to \mathbb{R}^k$ be a map. f is called differentiable of class C^r, iff a map $F : U \to \mathbb{R}^k$ with $A \subseteq U \subseteq_{open} \mathbb{R}^n$ exists that is differentiable of class C^r in the ordinary sense and that fulfills $F|_A = f$, i.e., F restricted to A is identical to f. F is called a differentiable C^r-extension of f.*

For example, $f : [0,1] \to \mathbb{R}^2, x \mapsto x^2$ is a C^∞-map with $F : \mathbb{R} \to \mathbb{R}^2, x \mapsto x^2$ being a differentiable C^∞-extension of f.

Based on this definition, we can define differentiable maps between manifolds (see also Figure 12):

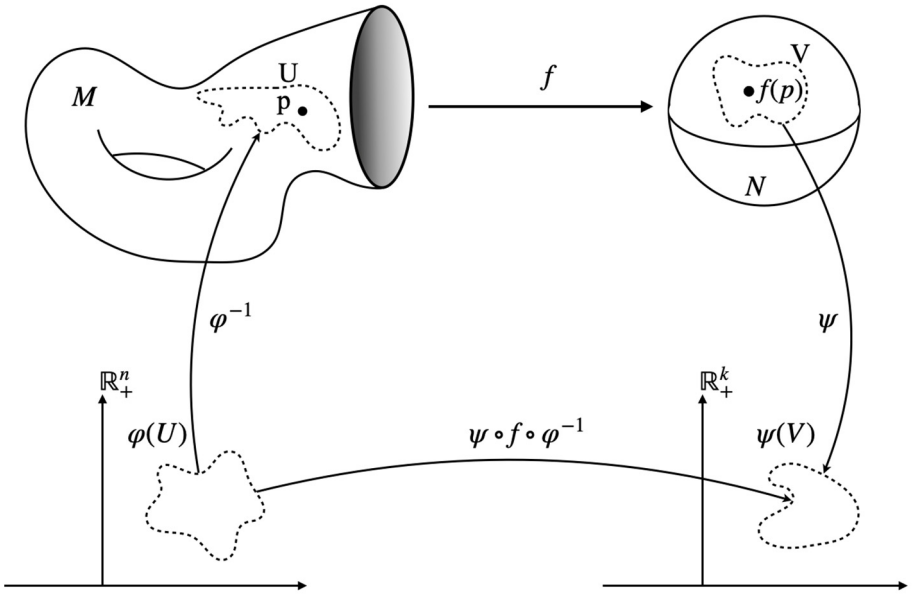


Figure 12. Differentiability of a Map Between Differentiable Manifolds.

Definition 6. *Let M and N be two C^r-manifolds (with or without boundary) with $dim M = n$ and $dim N = k$. A map $f : M \to N$ is said to be differentiable of class C^r, iff for every $p \in M$ there exist a chart (U, φ) of M around p and a chart (V, ψ) of N around $f(p)$ with $f(U) \subseteq V$, such that $\psi \circ f \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ is differentiable of class C^r.*

Recall that $\mathbb{R}^k$ is a manifold with the atlas $\mathfrak{A} = \left\{\left(\mathbb{R}^k, id\right)\right\}$. Thus, the definition before covers the definition of differentiability of maps $f : M \to \mathbb{R}^k$.

Maps that maintain the differentiable structure of manifolds are of special interest:

Definition 7. *Let M and N be two C^r-manifolds. A map $f : M \to N$ is said to be a C^r-diffeomorphism iff f is differentiable of class C^r, is bijective, and has an inverse that is also differentiable of class C^r. Furthermore, the manifolds M and N are called C^r-diffeomorphic.*

For example, $f : \mathbb{R} \to \mathbb{R}^2, x \mapsto x^3$ is a C^∞-map. According to Lemma 9, $N = \left\{(x, y) \mid y = x^3\right\}$ is a C^∞-manifold and so is $\mathbb{R}$ as a Euclidean space. $f : \mathbb{R} \to N$ is bijective with $g : N \to \mathbb{R}$, $(x, y) \mapsto x$ as inverse map, and g is also of class C^∞. Thus, $f : \mathbb{R} \to N$ is a C^∞-diffeomorphism.

Diffeomorphic manifolds cannot be distinguished based on differential topological properties. A diffeomorphism in differential topology plays the same role as homeomorphisms in general topology or isomorphisms in algebra.

Next, we show that a chart is a diffeomorphism; this is a well-known fact but the proof is instructive:

Note 6. *Let M be C^r-manifold (with or without boundary) with atlas $\mathfrak{A}$ and let $(U, \chi) \in \mathfrak{A}$ be a chart. Then, $\chi : U \to \chi(U)$ is a C^r-diffeomorphism.*

Proof. First, we have to show that χ is of class C^r; i.e., we have to show that $\psi \circ \chi \circ \varphi^{-1} : \varphi(W) \to \psi(V)$ is of class C^r for a chart (W, φ) of M and a chart (V, ψ) of $\mathbb{R}^k$. Choose $(W, \varphi) = (U, \chi)$ as a chart of M and $(V, \psi) = \left(\mathbb{R}^k, id\right)$ as a chart of $\mathbb{R}^k$. Then $\psi \circ \chi \circ \varphi^{-1} = id \circ \chi \circ \chi^{-1} = id : \chi(U) \to \mathbb{R}^k$, which is of class C^r.

χ is a homeomorphism, i.e., χ is bijective. Similar to before it is seen that χ^{-1} is of class C^r. □

The general question is about the relevance of the differentiability class C^r. A famous theorem by Whitney [44] proves that studying C^∞-manifolds suffice (a detailed proof of this theorem can be found in [37], 2–2.10). Because of this theorem, the restriction to smooth manifolds (i.e., C^∞-manifolds) is justified:

Theorem 1. *Let* $\mathbf{1} \leq \boldsymbol{r} \leq \infty$. *Then: Every* $\boldsymbol{C^r}$*-manifold is* $\boldsymbol{C^r}$*-diffeomorphic to a* $\boldsymbol{C^\infty}$*-manifold.*

4.5. Differential of a Map

A point $p \in M$ of a differentiable manifold M (with or without boundary) is associated with its *tangent space* T_pM (see Figure 13). This tangent space can be imagined as the set of all tangent vectors to M through p; the precise definition is much more subtle and complex (see [23], chapter 3) but for our purpose this descriptive idea suffices.

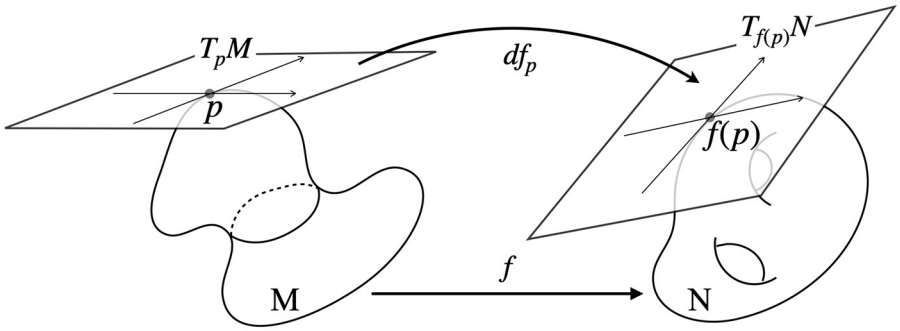


Figure 13. Tangent space and Differential.

T_pM is a vector space, and if M has dimension m it is $dim T_pM = m$ ([23] Proposition 3.12). The disjoint union of the tangent spaces of all points $p \in M$ is referred to as *tangent bundle* TM of M: $TM = \biguplus_{p\in M} T_pM$. With $dim M = m$, TM is a differentiable manifold with $dim TM = 2m$ ([23] Proposition 3.18).

If $f : M \to N$ is a differentiable map between two differentiable manifolds, then the *differential* df_p of f at $p \in M$ is a linear map $df_p : T_pM \to T_{f(p)}N$ (see Figure 13). As before, the precise definition is quite complex ([23], chapter 3), but, again, a vague intuition suffices for our purpose, especially because the differential df_p corresponds to the Jacobian matrix of f ([23], p. 61 ff). We will need the latter representation of the differential of a map because it allows us to compute the rank of the differential as the rank of the Jacobian matrix.

In case $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$ are "just" Euclidean spaces, M and N are smooth manifolds (see Note 5a), and their tangent spaces are the very same Euclidean spaces, i.e., $T_x\mathbb{R}^m = \mathbb{R}^m$ and $T_y\mathbb{R}^n = \mathbb{R}^n$ for any points $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ ([23] proposition 3.13). The differential $df_p : T_pM \to T_{f(p)}N$ of a map $f : M \to N$ becomes the total derivative $Df(p) = (\partial f_i(p)/\partial x_j)$ ([23] proposition C.3); thus, for $v \in T_pM$ it is $df_p(v) = Df(p)(v) = D_vf(p)$ where $D_vf(p)$ is the directional derivative of f in the direction v.

For example, $f : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x^2, y^3)$ is a C^∞-map. Its differential $df_{(x,y)} : \mathbb{R}^2 \to \mathbb{R}^2$ is $df_{(x,y)} = \begin{pmatrix} 2x & 0 \\ 0 & 3y^2 \end{pmatrix}$. For $(0, 1) \in T_{(x,y)}\mathbb{R}^2$ it is $df_{(x,y)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2x & 0 \\ 0 & 3y^2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3y^2 \end{pmatrix}$.

4.6. Maps of Constant Rank

For each linear map $L : V \to W$ between vector spaces V and W the rank of L is defined as the dimension of the image of L, i.e., $rank\ L := dim\ img\ L$. Note that in practice, the rank of a linear map is determined as the number of linear independent columns of the matrix representing the map L (see [34], Section 5.3, for more details). Since the image of L is a subspace of W, it is always $rank\ L \leq dim\ W$. According to the Dimension

Formula of Linear Algebra (see [34] Theorem 10.9) it is $rank\ L + dim\ ker L = dim\ V$; thus, $rank\ L = dim\ V - dim\ ker L \leq dim V$. This proves the following:

Note 7. *Let $L : V \to W$ be a linear map. Then: $rank\ L \leq min\{dim\ V, dim\ W\}$.*

The rank of differentiable maps $f : M \to N$ in $p \in M$ is the rank of the linear map df_p. It is key to the study of local and global properties of differentiable functions f.

Definition 8. *Let $f : M \to N$ be a differentiable map between the two differentiable manifolds M and N. The rank of f at $p \in M$ is the rank of its differential $df_p : T_pM \to T_{f(p)}N$, in symbols $rank_p f$. If f has the same rank r at every point $p \in M$ then f is said to have constant rank, in symbols $rank f = r$.*

According to the end of Section 4.5, the map $f : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x^2, y^3)$ has the differential $df_{(x,y)} = \begin{pmatrix} 2x & 0 \\ 0 & 3y^2 \end{pmatrix}$. With $df_{(x,y)} = \begin{pmatrix} 2x & 0 \\ 0 & 3y^2 \end{pmatrix}$ it is $df_{(0,0)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, i.e., it is $rank_{(0,0)} f = 0$. For $(x, y) \neq (0, 0)$ it is $2x \neq 0$ and $3y^2 \neq 0$, i.e., $rank_{(x,y)} f = 2$. And for $x = 0$ but $y \neq 0$ is $2x = 0$ and $3y^2 \neq 0$, i.e., $rank_{(0,y)} f = 1$. For $(x, y) \in U := \{(x, y) \mid x > 0 \wedge y > 0\} \subseteq_{open} \mathbb{R}^2$ is $rank_{(x,y)} f = 2$, i.e., f has a constant rank on U.

Since $df_p : T_pM \to T_{f(p)}N$ is linear, it is $rank_p f \leq min\{dimM, dimN\}$ (see Note 7 before). In case $rank_p f = min\{dimM, dimN\}$, f is said to have *full rank* at p and just full rank if it has full rank at every point of M. Maps of full rank have special names:

Definition 9. *Let $f : M \to N$ be a differentiable map between the two differentiable manifolds M and N. f is called an immersion in case $rank f = dim M$, i.e., df_p is injective. f is called a submersion in case $rank f = dim N$, i.e., df_p is surjective.*

In our context, immersions are of special interest. Thus, give some geometric intuition of an immersion, which is helpful in what follows. Remember that the differentiability of a map f at a point p means that it can be locally approximated by a linear map, i.e., by its differential df_p: for points x close to p it is $f(x) \approx f(p) + df_p \cdot x$. Thus, the properties of the differential approximate locally properties of the map. Injectivity of the linear map df_p means that it maintains the independence of "all directions" in T_pM when mapping to $T_{f(p)}N$: no two different vectors are "smashed" together. For the corresponding map f this translates into the fact that f does not "fold" or "collapse" parts of a neighborhood of the point p. Since an immersion f has a constant rank, which means that f transforms a neighborhood of each point p somehow "faithful", not introducing any "crushing".

An often-used well-known property is that the rank of a map can locally not decrease:

Lemma 4. *Let $f : M \to N$, $p \in M$ and $rank_p f = k$. Then there is a neighborhood $U \in \mathfrak{U}_p$ such that $rank_x f \geq k$ for each $x \in U$.*

Proof. It is $rank df_p = k$; thus, there is $k \times k$-submatrix $A(p)$ of df_p with $det A(p) \neq 0$. W.l.o.g. this submatrix is

$$A(p) = \left(\frac{\partial f_i}{\partial x_j}(p) \right)_{1 \leq i,j \leq k}$$

Next, define the map

$$\Delta : M \to \mathbb{R}\ ,\ x \mapsto det \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{1 \leq i,j \leq k}.$$

According to the Leibniz formula from linear algebra it is

$$\Delta(x) = detdf_x = \sum_{\sigma \in S_k} \left(sgn(\sigma) \prod_{i=1}^{k} \frac{\partial f_i}{\partial x_{\sigma(i)}}(x) \right),$$

i.e., Δ is continuous (even differentiable) because f is differentiable and the Leibniz formula is a polynomial. Thus, with $\Delta(p) = detA(p) \neq 0$ there is an $U \in \mathfrak{U}_p$ such that $\Delta(x) \neq 0$ for each $x \in U$. Consequently, $rank_x f \geq k$ for each $x \in U$. □

Although the rank of a map can locally not decrease, it may increase. For example, for $f(x,y) := (y, x^2 + y)$ it is $df_{(x,y)} = \begin{pmatrix} 0 & 2x \\ 1 & 1 \end{pmatrix}$. Thus, $df_{(0,0)} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ and $rank_{(0,0)} f = 1$. Arbitrary close to $(0,0)$ it is $df_{(\varepsilon,\varepsilon)} = \begin{pmatrix} 0 & 2\varepsilon \\ 1 & 1 \end{pmatrix}$, i.e., for $\varepsilon > 0$ it is $rank_{(\varepsilon,\varepsilon)} f = 2$. However, if the rank is already maximal, it is locally constant since it cannot decrease locally; this proves:

Lemma 5. *Let $f : M \to N$ be of full rank at $p \in M$. Then it is of full rank in a neighborhood $U \in \mathfrak{U}_p$.*

This has an important well-known implication:

Corollary 1. *Let M, N be two differentiable manifolds and $f : M \to N$ a differentiable map. Then:*

a. If df_p is surjective, then there exists a $U \in \mathfrak{U}_p$ such that $f|_U$ is a submersion.
b. If df_p is injective, then there exists a $U \in \mathfrak{U}_p$ such that $f|_U$ is an immersion.

Proof. (a) By assumption df_p is surjective, i.e., by definition $rank_p f = dimN$. But $rank_p f \leq min\{dimM, dimN\}$ implies that $dimN \leq min\{dimM, dimN\}$. Thus, $dimN = min\{dimM, dimN\}$ because $dimN < min\{dimM, dimN\}$ would be a contradiction. Consequently, $rank_p f = min\{dimM, dimN\}$ which shows that f is of full rank. The lemma before proves the claim. (b) is proven the same way. □

The next theorem (whose proof can be found in [23], Theorem 4.5) shows that in case the differential is bijective at a point of a manifold without boundary the map is a local diffeomorphism around that point. Note, that the precondition that the manifold must have no boundary is essential here: The inclusion $\iota : \mathbb{R}^n_+ \hookrightarrow \mathbb{R}^n$ is corresponding counter-example.

Theorem 2 (Inverse Function Theorem). *Let M, N be two differentiable manifolds without boundary and $f : M \to N$ a differentiable map. If df_p is bijective there are $U \in \mathfrak{U}_p$ and $V \in \mathfrak{U}_{f(p)}$ such that $f|_U : U \to V$ is a diffeomorphism.*

In the theorem before the conditions that both manifolds must have no boundary can be weakened: the codomain N may have a boundary but the image of f must be in the interior of the codomain, i.e., $f(M) \subseteq N \setminus \partial N$; the reason is that $N \setminus \partial N \subseteq_{open} N$ is a submanifold without boundary (see Note 5c), i.e., the original inverse function theorem applies.

Special kinds of immersions play an important role:

Definition 10. *Let $f : M \to N$ be a differentiable map between the two differentiable manifolds M and N. f is called an embedding iff f is an immersion and $f : M \to f(M)$ is a homeomorphism onto $f(M) \subseteq N$ in the subspace topology.*

The map $f : \mathbb{R} \to \mathbb{R}^2, x \mapsto (x^3, 0)$ is a homeomorphism, f is smooth, but because of $df_0 = 0$, f is not an immersion, thus, no embedding. Thus, not every smooth homeomorphism is automatically an embedding. However, under the following condition injective immersions are already embeddings (the proof can be found in [23], Proposition 4.22):

Lemma 6. *Let $f : M \to N$ be an injective immersion, M and N be manifolds with or without boundary. If any of the following holds, f is an embedding:*

a. M is compact.
b. M has no boundary, and $dim M = dim N$.

The following theorem is the basis of many other theorems in differential topology (see [45], Theorem 3.7.5):

Theorem 3 (Rank Theorem). *Let M, N be two differentiable manifolds and $f : M \to N$ a differentiable map. Let $p \in M$, $q := f(p) \in N$ and let $U \in \mathfrak{U}_p$ be a neighborhood of p such that for each $x \in U$ it is $rank_x f = r$ (i.e., f has constant rank in U). Then there is a chart for M around p and a chart for N around q, such that f has in these charts the form $f(x_1, \ldots, x_r, x_{r+1}, \ldots, x_m) = (x_1, \ldots, x_r, 0, \ldots, 0)$.*

Several key well-known properties of maps are inherited by their composition:

Note 8. *Let L, M, N be differentiable manifolds, and let $f : L \to M$ and $g : M \to N$ be maps. Then:*

a. If f and g are injective or surjective or bijective then $g \circ f$ is injective or surjective or bijective.
b. If f and g are immersions then $g \circ f$ is an immersion.
c. If f and g are continuous then $g \circ f$ is continuous.
d. If f and g are homeomorphisms then $g \circ f$ is a homeomorphism.
e. If f and g are embeddings then $g \circ f$ is an embedding.

Proof. (a) $g(f(x)) = g(f(y))$ implies $f(x) = f(y)$ because g is injective. Next, injectivity of f implies $x = y$. This proves the injectivity of $g \circ f$.

$f(L) = M$ and $g(M) = N$, thus, $(g \circ f)(L) = N$.This proves the surjectivity of $g \circ f$. Together, this proves the bijectivity of $g \circ f$.

(b) df and dg are injective. According to the chain rule it is $d(g \circ f) = dg \circ df$, i.e., part (a) shows that $d(g \circ f)$ is injective. Thus, $g \circ f$ is an immersion,

(c) Let $O \subseteq N$ be open. Because g is continuous, $g^{-1}(O) \subseteq M$ is open. Because f is continuous, $f^{-1}\left(g^{-1}(O)\right) \subseteq L$ is open. With $f^{-1} \circ g^{-1} = \left(g \circ f\right)^{-1}$ it follows that $\left(g \circ f\right)^{-1}(O) \subseteq L$ is open. Thus, $g \circ f$ is continuous.

(d) f and g are bijective, so is $g \circ f$ (see part (a)). f and g are continuous, so is $g \circ f$ (see part (c)). f^{-1} and g^{-1} are continuous, so is $\left(g \circ f\right)^{-1} = f^{-1} \circ g^{-1}$ (see part (c)). Thus, $g \circ f$ is a homeomorphism.

(e) f and g are embeddings, i.e., both are immersions as well as homeomorphisms. Because of part (b) $g \circ f$ is an immersion, and because of part (d) $g \circ f$ is a homeomorphism. Thus, $g \circ f$ is an embedding. □

4.7. Submanifolds

It is important to note that the definition of a manifold is completely independent of any surrounding space like a Euclidean space. In this sense, manifolds are abstract entities. Their concept has been introduced by Bernhard Riemann in 1854 (published 1868 [46]). It generalizes objects like curves and surfaces that had been studied before that time. The latter are entities within a Euclidian space. Thus, it is natural to ask whether any (abstract)

manifold is "equivalent" (also known as. "diffeomorphic") to a corresponding entity in a Euclidean space: this has been proven by Hassler Whitney in 1936 [44]. In this section, we summarize the corresponding concepts and results as relevant in our context.

A manifold maybe contained in another manifold (see Figure 14).

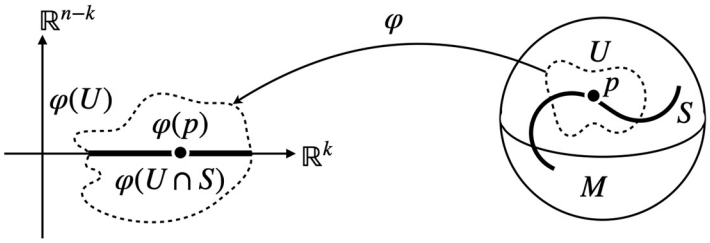


Figure 14. A Chart of an Embedded Submanifold.

Definition 11. *Let M be differentiable C^r-manifold, $dim M = n$ and let $S \subset M$ be a subset of M . S is called an embedded submanifold of M of dimension k (or of codimension $n-k$, respectively) and class C^r if for any point $p \in S$ there is a chart (U, φ) of M around p such that $\varphi(U \cap S) = \varphi(U) \cap \left(\mathbb{R}^k \times \{0\}\right)$.*

Sometimes, embedded submanifolds are also called *regular submanifolds*. Figure 14 depicts the situation. If $\mathfrak{A} = \{(U_i, \varphi_i) | i \in I\}$ is an atlas of M then $\mathfrak{A}|_S := \left\{\left(U_i \cap S, \varphi_i|_{U_i \cap S}\right) \middle| i \in I\right\}$ is an atlas of S. Note that the latter assumes that S is a topological subspace of M, i.e., the topology of S is the subspace topology. This will become important soon.

Especially, since $\mathfrak{A}|_S$ is an atlas of S, S is a C^r-manifold in itself:

Note 9. *Every embedded submanifold $\mathbf{S}$ of class $\mathbf{C^r}$ and dimension $\mathbf{k}$ is a $\mathbf{C^r}$-manifold with $\mathbf{dim S = k}$.*

The following lemma motivates the name "embedded" submanifold (see [47], Theorem 11.14 for a proof):

Lemma 7. *Let M be differentiable C^r-manifold, and let $S \subset M$ be an embedded submanifold of M. Then, the inclusion $\iota : S \hookrightarrow M$ is a C^r-embedding (and, thus, by definition an immersion).*

Vice versa, the name "embedding" of a map is justified by the following lemma (see [47], Theorem 11.13 for a proof):

Lemma 8. *Let M and S be differentiable C^r-manifolds, and let $f : S \to M$ be a C^r-embedding. Then, the image $f(S) \subseteq M$ is an embedded submanifold of M.*

A two-dimensional surface is an example of a manifold embedded in the three-dimensional Euclidian space. In general, graphs of differentiable functions are examples of such surfaces. The following lemma and definition generalizes the corresponding situation (see [23], Proposition 5.4):

Lemma and Definition 9. *Let M and N be C^r-manifolds (N with or without boundary), $dim M = m$ and $dim N = n$, $U \subseteq_{open} M$ and $f : U \to N$ be of class C^r. Then the graph $\Gamma(f) \subseteq M \times N$ is an embedded C^r-manifold of dimension m without boundary.*

Here, the graph of f is defined as $\Gamma(f) := \{(x, y) \in M \times N | x \in U \wedge y = f(x)\}$.

The Euclidean space $\mathbb{R}^n$ is a manifold, so its embedded submanifolds are of interest. The following is an often-used mechanism to produce embedded submanifolds in $\mathbb{R}^n$: For $U \subseteq_{open} \mathbb{R}^k$ and a C^r-differentiable map $f : U \to \mathbb{R}^m$ the graph $\Gamma(f) \subseteq \mathbb{R}^k \times \mathbb{R}^m$ is an

embedded C^r-manifold of dimension k without boundary (according to the lemma just before). If $f : U \subseteq \mathbb{R}^{n-1} \to \mathbb{R}$ is of class C^r, the graph $\Gamma(f)$ is a *hypersurface* in $\mathbb{R}^n$, i.e., a C^r-manifold of dimension $n-1$ (see Figure 15); obviously, this generalizes the notion of a surface in $\mathbb{R}^3$.

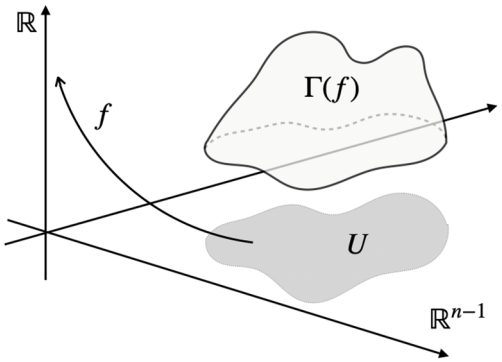


Figure 15. The Graph of a Map as Manifold.

Another important means to obtain embedded submanifolds is via so-called level-sets and regular values:

Definition 12. *Let $f : M \to N$ be a C^r-map.*

a. For any point $q \in N$ the set $f^{-1}(q)$ is called a level-set of f.

b. A point $p \in M$ is called a regular point if df_p is surjective (p is called a critical point otherwise).

c. $q \in N$ is called a regular value if each point of $f^{-1}(q)$ is a regular point (q is called critical value otherwise).

d. If $q \in N$ is a regular value the level-set $f^{-1}(q)$ is called a regular level-set.

Obviously, each point of M is a critical point if $dimM < dimN$ (df_p cannot be surjective in this case). Any regular level set is a C^r-manifold (see [23], Corollary 5.14):

Lemma 10. *Let M and N be C^r-manifolds, and $f : M \to N$ be of class C^r. Then, any regular level-set $f^{-1}(q) \subseteq M$ is an embedded submanifold of M with $dim f^{-1}(q) = dimM - dimN$.*

In Figure 16, $f^{-1}(c_1)$, $f^{-1}(c_2)$, and $f^{-1}(c_3)$ are three level-sets of the differentiable map f. Assuming that both c_1 and c_2 are regular values, the two level sets $f^{-1}(c_1)$ and $f^{-1}(c_2)$ are embedded submanifolds (the first one diffeomorphic to a circle, the second one diffeomorphic to two disjoint circles) of M of dimension 1 because of $dimM = 2$ and $dimN = 1$. But $f^{-1}(c_3)$ is not an embedded submanifold because it has the shape of an "8", i.e., it contains a singularity in the form of a self-intersection (see Section 4.3). This implies that c_3 is a critical value.

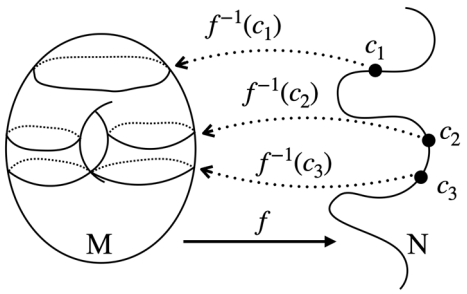


Figure 16. Level-Sets of a Differentiable Map.

The lemma before allows us to prove that the set of unitary transformations of a complex vector space is a manifold:

Lemma 11. *$\boldsymbol{U(n)}$ is a smooth compact connected manifold (even an embedded submanifold) with $\boldsymbol{dim_{\mathbb{R}}U(n) = n^2}$.$\boldsymbol{U(n)}$ is even a Lie-group.*

Proof. It is $U(n) = \{A \in GL(n,\mathbb{C}) \mid A^*A = I\}$, and $GL(n,\mathbb{C}) = \mathbb{C}^{n^2} = \mathbb{R}^{2n^2}$; thus, it is $U(n) \subset \mathbb{R}^{2n^2}$. Define $f : \mathbb{R}^{2n^2} = GL(n,\mathbb{C}) \to GL(n,\mathbb{C}) = \mathbb{R}^{2n^2}$ by $A \mapsto A^*A$; f is differentiable because building the conjugate transpose of a matrix is differentiable, and the multiplication of two matrices is differentiable too. The differential of f is $df_A(V) = A^*V + V^*A$ (see Appendix A).

Let $H(n) = \{A \in GL(n,\mathbb{C}) \mid A = A^*\}$ be the set of all Hermitian matrices. $H(n)$ is a $\mathbb{R}$-vector space of dimension n^2 ([34] Lemma 13.15). Thus, $H(n)$ is a smooth manifold of dimension n^2.

Because $f(B)^* = (B^*B)^* = B^*(B^*)^* = B^*B = f(B)$ for each $B \in GL(n,\mathbb{C})$, it is $f(GL(n,\mathbb{C})) \subseteq H(n)$, i.e., f is, in fact, a map $f : GL(n,\mathbb{C}) \to H(n)$ (and f is differentiable).

Next, we show that any unitary map $A \in U(n)$ is a regular point of $f : GL(n,\mathbb{C}) \to H(n)$, i.e., that the differential df_A is surjective for each $A \in U(n)$:

Choose an arbitrary $B \in H(n)$ and define $W := \frac{1}{2}AB \in GL(n,\mathbb{C})$. Then, $df_A(W) = \frac{1}{2}A^*AB + \frac{1}{2}B^*A^*A \overset{(1)}{=} \frac{1}{2}IB + \frac{1}{2}B^*I = \frac{1}{2}B + \frac{1}{2}B^* \overset{(2)}{=} \frac{1}{2}B + \frac{1}{2}B = B$ (hereby, (1) is because $A^*A = I$ and (2) because $B^* = B$).

Now it is $f^{-1}(I) = U(n)$; thus, $I \in H(n)$ is a regular value and $U(n) = f^{-1}(I)$ is a regular level set. Consequently, $U(n)$ is an embedded submanifold (Lemma 10), and, thus, a manifold itself (Note 9); furthermore, $dim U(n) = dim f^{-1}(I) = dim GL(n,\mathbb{C}) - dim H(n) = 2n^2 - n^2 = n^2$.

Next, we prove compactness.

(a) $f : GL(n,\mathbb{C}) \to GL(n,\mathbb{C})$ with $A \mapsto A^*A$ is especially continuous. Now, $U(n) = f^{-1}(I)$, $\{I\} \subset GL(n,\mathbb{C})$ is a closed set, and pre-images of closed sets under continuous maps are closed, i.e., $U(n) \subset_{closed} \mathbb{R}^{2n^2}$.

(b) Next, $A^*A = I$ means especially that the columns of $A = (a_{ij})$ are unit vectors, i.e., $\sum_i \left|a_{ij}\right|^2 = 1$ for each $1 \le j \le n$. Thus, $\left|a_{ij}\right|^2 \le 1$ for all $1 \le i,j \le n$. This implies that $\| A \|^2 = \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 \le n^2$ for each $A \in U(n)$, i.e., $U(n)$ is bounded in $\mathbb{R}^{2n^2}$. Thus, according to the theorem of Heine-Borel, $U(n)$ is compact in $\mathbb{R}^{2n^2} = GL(n,\mathbb{C})$.

Finally, we prove connectedness.

Le $A \in U(n)$. Then, A is diagonalizable, i.e., there exists a unitary matrix T such that $A = T diag\left(e^{i\varphi_1}, \cdots, e^{i\varphi_n}\right)T^*$ ([34] Theorem 18.13). Let $0 \le t \le 1$; then, $A_t := T diag\left(e^{it\varphi_1}, \cdots, e^{it\varphi_n}\right)T^* \in U(n)$ (because $A, B \in U(n) \Rightarrow AB \in U(n)$). It is $A_0 = I$ and $A_1 = A$, i.e., there is a path from the identity matrix I to A. Thus, any two unitary matrices can be connected by a path (e.g., via I): $U(n)$ is path-connected. Since every path-connected topological space is connected ([35] Proposition 9.26), $U(n)$ is connected.

$U(n)$ is a Lie group because matrix multiplication is a differentiable map $\cdot : U(n) \times U(n) \to U(n)$, $(a_{ij}) \cdot (b_{st}) = \left(\sum_j a_{ij}b_{jt}\right)_{it}$, and for two unitary matrices $A, B \in U(n)$ it is $(AB)(AB)^* = ABB^*A^* = AIA^* = AA^* = I$, i.e., $AB \in U(n)$. □

Note that the condition "regular" is key in Lemma 10. Without that condition, any closed subset of a manifold can be made the level set of a differentiable function (see [23], Theorem 2.29 for a proof):

Lemma 12. *Let M be a differentiable C^r-manifold, and let $K \subseteq_{closed} M$ be a closed subset of M. Then there exists a C^r-function $f : M \to \mathbb{R}$ with $f^{-1}(0) = K$.*

In many practical situations, both, M and N are Euclidean spaces. For example, with $M = \mathbb{R}^{n+1}$ and $N = \mathbb{R}$, the map $f : \mathbb{R}^{n+1} \to \mathbb{R}$, $x \mapsto |x|^2$ has the differential $df_x = 2(x_1 \cdots x_n) \neq 0$ for $x \neq 0$, i.e., $rank_x f = 1$ for $x \neq 0$; thus, df_x is surjective, except at the origin. This implies that each $c \neq 0 \in \mathbb{R}$ is a regular value; according to the lemma before, $f^{-1}(c)$ is a regular level-set for each $c \neq 0$. Thus, $f^{-1}(c) = \left\{ x \in \mathbb{R}^{n+1} \mid |x|^2 = c \right\}$ is a sphere of dimension n, especially the unit sphere $\mathbb{S}^n = f^{-1}(1)$ is an n-dimensional embedded submanifold of $\mathbb{R}^{n+1}$.

There are important situations (e.g., in the context of Lie groups—see [48]) in which the notion of a submanifold has to be generalized. Very roughly, any differentiable manifold that is a subset of another differentiable manifold and is "properly situated" there is considered a certain kind of a submanifold. More precisely,

Definition 13. *Let M and N be C^r-manifolds and $f : N \to M$ be an injective immersion of class C^r. Then, the image $f(N) \subseteq M$ is called an immersed submanifold of M.*

For example, let $f :]-\pi, +\pi[\to \mathbb{R}^2$, $x \mapsto (sin2x, sinx)$. Then, f is injective on the open interval $]-\pi, +\pi[$ (see Appendix B for more details). Also, f is an immersion because $df_x = (2cos2x, cosx) \neq 0$, i.e., $rank_x f = 1$ for $x \in]-\pi, +\pi[$. Thus, f is an injective C^r-immersion, i.e., $S := f(]-\pi, +\pi[) \subseteq \mathbb{R}^2$ is an immersed submanifold (see Figure 17). Furthermore, $f :]-\pi, +\pi[\to S$ is also surjective, thus, bijective, but f is not a homeomorphism because the image $S \subseteq_{compact} \mathbb{R}^2$ is compact with the subspace topology (it is closed and bounded), $]-\pi, +\pi[\subseteq \mathbb{R}$ is not compact, while compactness is a topological invariant. Together, $S \subseteq \mathbb{R}^2$ is an immersed submanifold, but S is not an embedded submanifold of the manifold $\mathbb{R}^2$. Another argument supporting the latter: any neighborhood of the origin of S contains a singularity, namely a shape like in Figure 11b.

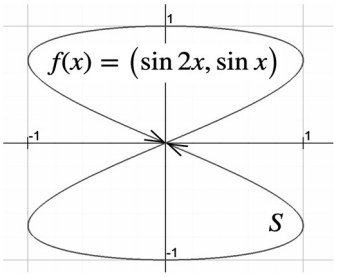


Figure 17. An Injective Immersion That is No Topological Embedding.

According to Lemma 7, for any embedded submanifold $S \subset M$ the inclusion $\iota : S \hookrightarrow M$ is an embedding and, thus, an injective immersion, i.e., $\iota(S) = S \subset M$ is an immersed submanifold of M. This proves the following:

Lemma 13. *Let M be a C^r-manifold (with or without a boundary), $S \subset M$ an embedded submanifold. Then, S is an immersed submanifold.*

The opposite is not true: the example before shows that an immersed submanifold is, in general, not an embedded submanifold. However, the following lemma ([23], Proposition 5.21) gives two situations in which an immersed submanifold is already an embedded submanifold:

Lemma 14. *Let M be a C^r-manifold (with or without a boundary), $S \subseteq M$ an immersed submanifold. If $dim S = dim M$ or if $S \subseteq_{compact} M$ then S is embedded.*

Finally, we answer the question posed at the beginning of this section: Any compact (abstract) manifold is diffeomorphic to an embedded submanifold of a Euclidean space of high enough dimension—more precisely,

Theorem 4. *Let $r \geq 1$ and let M be a compact C^r-manifold with boundary. M can be embedded into $\mathbb{R}^{2m+1}_+$ with $\partial M \subseteq \partial \mathbb{R}^{2m+1}_+ = \{x \in \mathbb{R}^{2m+1} | x_{2m+1} = 0\}$.*

Again, this theorem is by Whitney [44], and a detailed proof can be found in [37] (Theorem 1–4.3).

In summary, this section provided the necessary background of dimensional expressivity as needed in Section 5. Before presenting the corresponding details, the next section finally focuses on the unitary approach, giving a precise definition of the corresponding notion of expressivity. This requires explaining first how volumes on manifolds can be measured.

4.8. Volumes of Manifolds and the "Uniform Approach"

Section 2.3.1 motivated to define the expressivity of a variational quantum circuit by the "volume" of $\Omega_{\mathcal{C}}(\mathbb{P})$ in the unitary group $U(n)$: vividly, the larger this volume, the higher the likelihood that $\Omega_{\mathcal{C}}(\mathbb{P})$ hits the solutions $\mathcal{S}(X)$ of a given problem X.

4.8.1. Linear Approximations

Here, we provide more details about the notion of "the volume" of a differentiable manifold and especially how the notion of "volume" is related to the unitary group. For this purpose, we first remember that a function f is called differentiable at a point x if it can be approximated locally by a linear function:

$$f(x + \xi) = f(x) + L\xi + o(\| \xi \|) \tag{6}$$

Here, ξ is a point in a small neighborhood of x, L is a linear function, and $o(\| \xi \|)$ is the small error made when considering $f(x) + L\xi$ as the value of $f(x + \xi)$: in a small neighborhood, the differentiable function f is nearly the linear function L. For a differentiable function $f : \mathbb{R} \to \mathbb{R}$ this means that L is a real number and $f(x) + L\xi$ is the tangent at x at the graph of f; this tangent locally approximates the function f. Thus, the graph of f, i.e., the manifold $\Gamma(f)$ (see Lemma 9), is approximated by this tangent around $(x, f(x))$.

The idea of linear approximation can be used in arbitrary dimensions: in part (a) of Figure 18, a basis v_1, v_2 has been chosen for the tangent space T_pM of the manifold M at point p. The basis spans a parallelepiped $Q := \{a_1v_1 + a_2v_2 | 1 \leq a_1, a_2 \leq 1\}$. If we take small vectors v_1, v_2 the parallelepiped approximates the manifold M around p with a small error, i.e., the manifold looks locally like a very small parallelepiped of the tangent space.

Applying this linear approximation to enough points $p_1, \ldots, p_k$ of the manifold, i.e., if the manifold is "covered" by parallelepipeds, the manifold is turned into a linear approximation of the whole manifold: part (b) of Figure 18 depicts this for the upper part of a manifold which looks like small parallelepipeds glued together (in the direction of the tangent spaces).

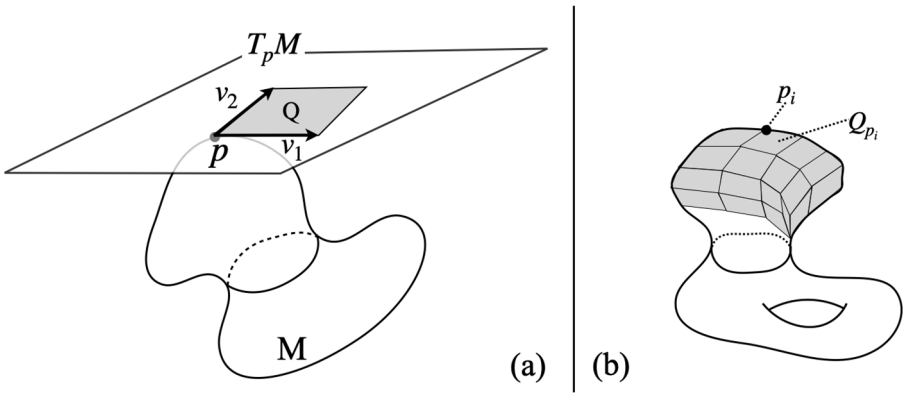


Figure 18. Linear Approximation of a Differentiable Manifold.

4.8.2. Approximating Volumes

Such a linear approximation allows us to compute the approximate volume of the manifold $vol(M)$ by computing the volume of the parallelepipeds $vol(Q_{p_i})$ and summing up their volumes:

$$vol(M) \approx \sum_{i=1}^{k} vol(Q_{p_i}) \tag{7}$$

The volume of a parallelepiped is computed as follows: let $v_1, \ldots, v_n$ be linear independent vectors, and let $Q := \{\sum_i a_i v_i | 1 \leq a_1, a_2, \ldots, a_n \leq 1\}$ be the parallelepiped spanned by these vectors. Then, from linear algebra, it is known that the volume $vol(Q)$ of the parallelepiped is

$$vol(Q) = |det(v_1 v_2 \ldots v_n)| \tag{8}$$

Define $V := (v_1 v_2 \ldots v_n)$ to be the matrix with columns v_1, v_2,...,v_n. Then:

$$\begin{aligned} vol(Q)^2 = |det(V)|^2 \quad &= det(V) \cdot det(V) \\ &= det\big(V^T\big) \cdot det(V) \\ &= det\big(V^T \cdot V\big) \\ &= det\big(\langle v_i, v_j \rangle\big)_{1 \leq i,j \leq n} \end{aligned}$$

Thus,

$$vol(Q) = \sqrt{det\big(\langle v_i, v_j \rangle\big)_{1 \leq i,j \leq n}} \tag{9}$$

4.8.3. Riemannian Manifolds and Volume Forms

Equation (9) reveals that computing the volume of a parallelepiped depends on a scalar product. Consequently, we need a scalar product for every tangent space of the manifold M.

Definition 14. *Let M be a differentiable manifold, and let g be a function that associates with each $p \in M$ a scalar product $g_p : T_pM \times T_pM \to \mathbb{R}$ "in a differentiable manner". Then, g is called a Riemannian metric on M, and (M, g) is called a Riemannian manifold.*

Note, that the phrase "in a differentiable manner" is left vague: a precise definition would require to define differentiable vector fields which we do not need in this paper. Also, the differentiability of g is not relevant in our context.

Consequently, we assume that M is a Riemannian manifold; in fact, every differentiable manifold is a Riemannian manifold ([39], Proposition 2.4). Then, the volume of a parallelepiped Q_p in T_pM is as follows:

$$vol(Q_p) = \sqrt{det\left(g_p(v_i, v_j)\right)_{1\leq i,j\leq n}} \tag{10}$$

With Equations (7) and (10) we can approximate the volume of a Riemannian manifold M as follows:

$$vol(M) \approx \sum_{i=1}^{k} vol(Q_{p_i}) = \sum_{i=1}^{k} \sqrt{det\left(g_p(v_i, v_j)\right)_{1\leq i,j\leq n}} \tag{11}$$

By choosing infinitesimal small parallelepipeds and correspondingly more and more points from the manifold we perform a limit process. In analogy to the limit process that defines the Riemannian integral we write very informally (and only conceptually) with $G = \left(g_p(v_i, v_j)\right)_{i,j}$:

$$vol(M) = \int_M \sqrt{det(G)}dx \tag{12}$$

The precise definitions behind this notion need a lot more concepts and machinery. The most fundamental concept needed is that of a volume form dV that abstracts our informal notation $\sqrt{det(G)}dx$. In our context, the unitary group admits such a volume form:

Note 10. *The unitary group $U(n)$ admits a (unique) volume form dV.*

Proof. Every differentiable manifold is a Riemannian manifold ([39], Proposition 2.4). If a Riemannian manifold M is oriented then it admits a (unique) volume form dV ([39], Proposition 2.41). According to Lemma 11, $U(n)$ is a differentiable manifold, thus, it is a Riemannian manifold also. Any Lie group is orientable ([49], Lemma 6). Since $U(n)$ is a Lie group (Lemma 13) it admits a (unique) volume form. □

The volume form dV of a manifold is (in local coordinates) $dV = \sqrt{det(G)}dx$; it allows (as Equation (12) indicates) us to compute the volume of M, namely $vol(M) = \int_M dV$. It also allows us to compute the integral of functions f, i.e., $\int_M fdV$ ([see 39], the discussion following Proposition 2.41).

4.8.4. Haar Measure

Computing volumes is tight to differentiable manifolds, not applicable to other "spaces". For this purpose, the concept of a *measure* is introduced (see [50] for details) that is more abstract than a volume but mimics its properties. Luckily, in our context, both concepts are the same. This is roughly seen as follows:

Each open set W of a differentiable manifold M is a differentiable manifold by itself (Note 5). Thus, the inclusion map $\iota : W \hookrightarrow M$ induces a Riemannian metric on W, and $vol(W) = \int_W dV$ is defined. This in turn, defines a measure on the Borel sets $\mathcal{B}(M)$ of M, where $\mathcal{B}(M)$ is the smallest σ-algebra containing all open sets of M (refer to [50] for details about Borel sets, measure spaces, and measures). This turns the manifold M into a measure space $(M, \mathcal{B}(M), vol)$ ([51], Section 1.7). A measure is more general than computing volumes via integrals.

As a Lie group, $U(n)$ admits a left-invariant measure, the so-called *Haar measure.* This measure is unique up to multiplication by a positive constant ([48], Theorem 3.1). A compact Lie group admits a bi-invariant Riemannian metric ([52], Proposition 2.17), which induces a bi-invariant volume form. Since the Haar measure is unique, we obtain the following:

Note 11. *The volume form dV on a compact Lie group is the Haar measure.*

Especially, since $U(n)$ is a Lie group, instead of abstractly speaking of the measure of a subset of a manifold, we can speak about its volume.

4.8.5. Volume of Submanifold

Having the ability to determine the volume of manifolds, the question of volumes of submanifolds comes up.

Note 12. *Let M be a manifold with volume element dV, and let $N \subseteq M$ be an (embedded or immersed) submanifold. If $dimN < dimM$ then $vol(N) = 0$.*

Proof. Let M, N be differentiable manifolds and let $f : N \to M$ be differentiable map. If $dimN < dimM$ then $f(N) \subset M$ has measure zero ([23], Corollary 6.11). For an (embedded or immersed) submanifold N of M the inclusion map $\iota : N \hookrightarrow M$ is differentiable (Lemma 7), i.e., if $dimN < dimM$ then N has measure zero in M. □

This implies that any submanifold S of the unitary group $U(n)$ with $dimS < dimU(n)$ has measure zero. Since measure and volume coincide it is $vol(S) = 0$ for $dimS < dimU(n)$:

Note 13. *Let $S \subset U(n)$ be a submanifold, $dimS < dimU(n)$. Then $vol(S) = 0$.*

Thus, $vol(\Omega_{\mathcal{C}}(\mathbb{P})) = 0$ in case $dim\Omega_{\mathcal{C}}(\mathbb{P}) < dimU(n)$ (if $\Omega_{\mathcal{C}}(\mathbb{P})$ is a submanifold at all—see next chapter). Consequently, the illustrative motivation about measuring expressivity in the "Unitary Approach" (Section 2.3.1) requires refinement.

4.8.6. Uniform Distribution

Let $A =]0, \varepsilon[\times]0, \varepsilon[\subset \mathbb{R}^2$ be a small open square in the Euclidian plain. If a point has to be picked randomly from A, the probability for any two points $x, y \in A$ being picked is the same: picking is "uniformly random". This means that the probability of a certain point being picked is identical to the probability of any other point in A being picked: the corresponding distribution of probabilities is uniform.

Obviously, the probability of a point being picked from A is related to the volume of A: the smaller A the higher the likelihood of a point within A to be picked. If we move the square around in the Euclidian plain, e.g., translating it by a vector $v \in \mathbb{R}^2$ to $A_v = v + A := \{v + t | t \in]0, \varepsilon[\times]0, \varepsilon[\} \subset \mathbb{R}^2$, the probability distribution remains the same because the volume is unchanged. This is based on the left-invariance of the volume in Euclidean space: by adding $v \in \mathbb{R}^2$ from left to each vector in A does not change the volume, i.e., $vol(A) = vol(A_v) = vol(v + A)$.

This is different for arbitrary manifolds: the Euclidean space if flat, but picking points from areas on curved manifolds may behave differently. For example, let $\mathbb{S}^2$ be the unit sphere. Then

$$\begin{array}{ccc} f :]0, \pi[\times]0, 2\pi[& \to \mathbb{S}^2 \\ (\vartheta, \varphi) & \mapsto (sin\vartheta cos\varphi, sin\vartheta sin\varphi, cos\vartheta) \end{array} \tag{13}$$

is a chart of $\mathbb{S}^2$ (leaving out both poles as well as the meridian of the sphere—otherwise f would not be a diffeomorphism). Figure 19 depicts the images under f of different parts of the domain:

- Part (a) shows the image of the whole domain, i.e., the sphere $\mathbb{S}^2$;
- Part (b) is the image of $]\pi/2 - \varepsilon, \pi/2 + \varepsilon[\times]0, 2\pi[$ (ε is a small positive number), resulting in a belt around the equator;
- Part (c) is the image of $]\pi - \varepsilon, \pi[\times]0, 2\pi[$ resulting in a cap of the north pole;
- Part (c) is the image of $]0, \varepsilon[\times]0, 2\pi[$ resulting in a cap of the south pole.

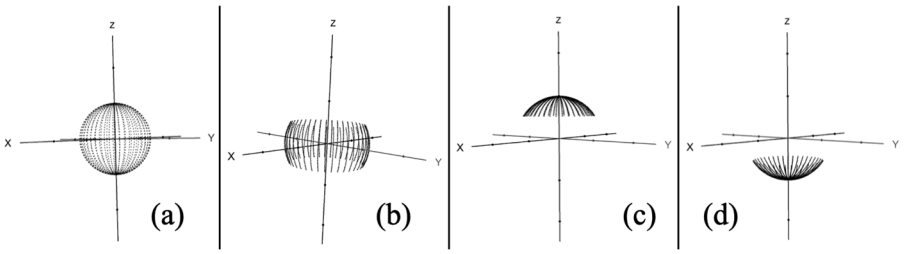


Figure 19. Volumes indicating non-uniform distributions.

The figure indicates that the volume of the belt is larger than the volume of a cap. But a larger volume of an area means a smaller probability of a point of the area being randomly picked. Thus, points on the sphere with values of ϑ close to $\pi/2$ (i.e., points from the belt) have a smaller probability of being randomly picked than points with values of ϑ close to π (i.e., points from the cap of the north pole) or close to 0 (i.e., points from the cap of the south pole): by picking a random latitude, points close to the poles have a higher probability to be picked than points near the equator. In this sense, points on the sphere are somehow "concentrated" towards the poles. As a consequence, the corresponding probability distribution is not uniform.

The volume that represents the probability distribution of randomly picking a point must reflect this effect of concentration to become a uniform distribution. As the figure indicates, shifting the belt to a pole does change its volume, i.e., the "usual" volume on the sphere is not translation invariant. But the Haar measure is translation invariant by definition. Thus, if we take the Haar measure to compute the probability distribution (i.e., the volume of areas of the sphere), the distribution becomes uniform: every area of the sphere of the same Haar measure has equal probability of containing a particular chosen point.

Consequently, the difference between the Haar measure of an area and its "usual" volume is an indicator of how much a distribution based on the usual volume deviates from being uniform.

4.8.7. Measuring Expressivity

"Deviation" can be assessed by various means. Especially, whether "volume" is computed directly or indirectly may differ. An example of an "indirect approach" is described next and is based on [6].

Whenever $A \in U(n)$ has been uniformly randomly chosen (often called "Haar random"), $A|0\rangle$ is a uniformly random state. Thus, if $S \subseteq U(n)$ is a set of Haar random unitary matrices, then $S|0\rangle := \{A|0\rangle \mid A \in S\}$ is a Haar random set of states. But $\Omega_{\mathcal{C}}(\mathbb{P}) \subset U(n)$ is not necessary Haar random, thus, the set $\Omega_{\mathcal{C}}(\mathbb{P})\,|0\rangle$ is not necessarily a set of uniformly random states.

Different approaches have been defined to compare $S|0\rangle$ and $\Omega_{\mathcal{C}}(\mathbb{P})\,|0\rangle$. One approach (see [4,6]) is as follows: first, the elements of these sets are turned into matrices, i.e., for $|\psi\rangle \in S|0\rangle$ and $|\varphi\rangle \in \Omega_{\mathcal{C}}(\mathbb{P})\,|0\rangle$ the density matrices $|\psi\rangle\langle\psi|$ and $|\varphi\rangle\langle\varphi|$ are taken. Note, that in fact $|\varphi\rangle = |\varphi_p\rangle$ depends on parameters $p \in \mathbb{P}$. For $S = U(n)$, the matrix $\int_{U(n)} \left|\psi\right\rangle\langle\psi| d\mu$ is computed as well as the matrix $\int_{\mathbb{P}} |\varphi_p\rangle\langle\varphi_p| dx$. Finally, the deviation between these two matrices is assessed and taken as expressivity of the ansatz $\mathcal{C}$. For example, based on a matrix norm, the expressivity $\eta(\mathcal{C})$ becomes

$$\eta(\mathcal{C}) := \| \int_{U(n)} |\psi\rangle\langle\psi| d\mu - \int_{\mathbb{P}} |\varphi_p\rangle\langle\varphi_p| dx \, \| \tag{14}$$

The smaller $\eta(\mathcal{C})$, the closer $\Omega_{\mathcal{C}}(\mathbb{P}) \subset U(n)$ becomes to being Haar random. If $\mathcal{C}$ and $\widehat{\mathcal{C}}$ are two ansatzes, ansatz $\mathcal{C}$ is called more expressive than ansatz $\widehat{\mathcal{C}}$ iff $\eta(\mathcal{C}) < \eta\left(\widehat{\mathcal{C}}\right)$. [31] gives a procedure how expressivity of a specific ansatz can be experimentally determined.

5. Circuit Manifolds

In this section we continue the discussion of dimensional expressivity that was interrupted by the section before giving the precise definition of expressivity according to the unitary approach. We provide a detailed proof that a state map (i.e., the map defined by a parameterized quantum circuit according to Definition 1) maps a properly chosen parameter space into a submanifold of $\mathbb{S}^{n-1}$.

5.1. State Maps Induce Local Immersions

As introduced in Section 2.3.2, let $\mathbb{P} \subseteq \mathbb{R}^k$ be a parameter space and $\mathcal{C}$ be an ansatz depending on the parameters $p_1, \ldots, p_k \in \mathbb{P}$, and let $\Lambda_{\mathcal{C}}$ be the map $\Lambda_{\mathcal{C}} : \mathbb{P} \to \mathbb{S}^{n-1}$ with $p_1, \ldots, p_k \mapsto \mathcal{C}(p_1, \ldots, p_k)\,|\iota\rangle$ with a chosen fixed initial state $|\iota\rangle$. Thus, $\Lambda_{\mathcal{C}}(\mathbb{P}) \subseteq \mathbb{S}^{n-1}$ is the set of all states reachable by the parameterized circuit $\mathcal{C}$. Lemma 1 shows that $\Lambda_{\mathcal{C}}$ is smooth, i.e., of class C^s for any $s \in \mathbb{N}$.

Assumption 1. *From now on, we assume that the parameter space $\mathbb{P} \subseteq \mathbb{R}^k$ is an embedded submanifold of $\mathbb{R}^k$ of dimension k. As discussed in* Section 3.2 *and* Section 3.3*, this is often the case in practical situations.*

Let $q \in \mathbb{P}$ and let $rank_q \Lambda_{\mathcal{C}} = r$. Thus, there is an $r \times r$ submatrix of the differential $d(\Lambda_{\mathcal{C}})_q$ whose determinant is not 0; w.l.o.g. with $d\left(\Lambda_{\mathcal{C}}\right)_q =: \left(\lambda_{ij}\right)_{1 \leq i,j \leq k}$ this submatrix is $(\lambda_{ij})_{1 \leq i,j \leq r}$. According to Lemma 4, there exists a neighborhood $W \in \mathfrak{U}_q$ open in $\mathbb{R}^k$ such that $rank_x \Lambda_{\mathcal{C}} \geq r$ for each $x \in W$; w.l.o.g. W is an open ball centered around q contained in $\mathbb{P}$. Next, let $\mathbb{R}^r_q := \left\{ x \in \mathbb{R}^k \mid x_j = q_j, j > r \right\}$ be the hyperplane of $\mathbb{R}^k$ parallel to $\mathbb{R}^r$ through q, and let $\widehat{\mathbb{P}} := \mathbb{R}^r_q \cap W$ (see Figure 20); i.e., $\widehat{\mathbb{P}}$ is the intersection of W with $\mathbb{R}^r_q$. If $W \subseteq \mathbb{P}$ is chosen to be small enough, it is contained in a chart (U, φ) of $\mathbb{P}$ around q, i.e., $W \subseteq U$. With $\widehat{\mathbb{P}} = \widehat{\mathbb{P}} \cap W$ open in $\widehat{\mathbb{P}}$ in the subspace topology, $\left\{ \left(\widehat{\mathbb{P}}, \varphi|_{\widehat{\mathbb{P}}}\right) \right\}$ is an atlas of $\widehat{\mathbb{P}}$. This shows the following:

Note 14. *$\widehat{\mathbb{P}} \subseteq \mathbb{P}$ is an embedded submanifold of $\mathbb{P}$ with $\boldsymbol{dim}\widehat{\mathbb{P}} = \boldsymbol{r}$.*

It is $\Lambda_{\mathcal{C}}(x_1, \ldots, x_r, \ldots, x_j, \ldots, x_k) = \Lambda_{\mathcal{C}}(x_1, \ldots, x_r, \ldots, \overline{x}_j, \ldots, x_k)$ for $j > r$ and for corresponding points in $\widehat{\mathbb{P}}$ (i.e., $x_j = q_j = \overline{x}_j$), thus, $\Lambda_{\mathcal{C}}$ is constant in each of the x_j directions. Consequently, $\partial \Lambda_{\mathcal{C}} / \partial x_j = 0$ for $j > r$. Together with $rank_x \Lambda_{\mathcal{C}} \geq r$ this implies $rank_x \Lambda_{\mathcal{C}} = r = dim\widehat{\mathbb{P}}$ for $x \in \widehat{\mathbb{P}}$. This proves the following:

Lemma 15. *$\widehat{\boldsymbol{\Lambda}}_{\boldsymbol{\mathcal{C}}} := \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{C}}}|_{\widehat{\mathbb{P}}} : \widehat{\mathbb{P}} \to \mathbb{S}^{\boldsymbol{n}-1}$ is an immersion.*

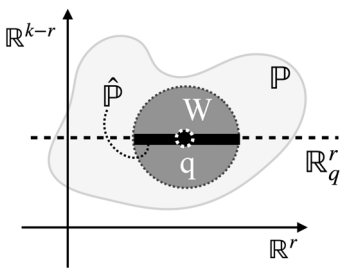


Figure 20. Restricting the Parameter Space.

5.2. Locally Embedded Circuit Manifolds

According to the Rank Theorem (Theorem 3), there exist charts around q and $f(q)$ such that $\widehat{\Lambda}_{\mathcal{C}}$ has the form $\widehat{\Lambda}_{\mathcal{C}}(x_1,,\ldots,x_k) = (x_1,\ldots,x_r,0,\ldots,0)$. I.e. $\widehat{\Lambda}_{\mathcal{C}}$ is injective in a neighborhood of $q \in \widehat{\mathbb{P}}$. Thus, for W chosen properly (i.e., by possibly shrinking it), $\widehat{\Lambda}_{\mathcal{C}}$ is injective, and with $\widehat{\Lambda}_{\mathcal{C}}$ being also an immersion, we achieve the following (Definition 13):

Lemma 16. $\widehat{\mathbf{\Lambda}}_{\boldsymbol{\mathcal{C}}} = \mathbf{\Lambda}_{\boldsymbol{\mathcal{C}}}|_{\widehat{\mathbb{P}}}$ *is an injective immersion, and* $\widehat{\mathbf{\Lambda}}_{\boldsymbol{\mathcal{C}}}\left(\widehat{\mathbb{P}}\right) \subseteq \mathbb{S}^{\boldsymbol{n}-1}$ *is an immersed submanifold.*

According to Note 14, $\widehat{\mathbb{P}} \subseteq \mathbb{P}$ is an embedded submanifold of $\mathbb{P}$ with $dim\widehat{\mathbb{P}} = r$. By Assumption 1, $\mathbb{P} \subseteq \mathbb{R}^k$ is an embedded submanifold of $\mathbb{R}^k$. Thus (Lemma 7), both inclusions $\widehat{\mathbb{P}} \hookrightarrow \mathbb{P}$ as well as $\mathbb{P} \hookrightarrow \mathbb{R}^k$ are embeddings, which implies that the composed inclusion $\iota : \widehat{\mathbb{P}} \hookrightarrow \mathbb{R}^k$ is an embedding (Note 8e). According to Lemma 8, the image of the inclusion $\iota\left(\widehat{\mathbb{P}}\right) = \widehat{\mathbb{P}} \subseteq \mathbb{R}^k$ is an embedded submanifold. This proves the following:

Note 15. $\widehat{\mathbb{P}} \subseteq \mathbb{R}^{\boldsymbol{k}}$ *is an embedded submanifold with* $dim\widehat{\mathbb{P}} = r$.

With $\widehat{\mathbb{P}} \subseteq \mathbb{R}^k$ being an embedded submanifold, choose a chart (U, φ) of $\widehat{\mathbb{P}}$ around q, i.e., it is $\varphi(U) \subseteq_{open} \mathbb{R}^r$. Next, chose a compact ball $B_\varepsilon(\varphi(q)) \subset_{compact} \varphi(U)$ with center $\varphi(q)$. Now, $B_\varepsilon(\varphi(q))$ is a compact manifold (with boundary) of dimension r, and φ^{-1} is a diffeomorphism; thus, $M := \varphi^{-1}(B_\varepsilon(\varphi(q))) \subset \widehat{\mathbb{P}}$ is a compact submanifold of dimension r. According to Lemma 16, $\widehat{\Lambda}_{\mathcal{C}}$ is an injective immersion, which implies that especially $\widehat{\Lambda}_{\mathcal{C}}|_M : M \to \mathbb{S}^{n-1}$ is an injective immersion. Then, according to Lemma 6a, $\widehat{\Lambda}_{\mathcal{C}}|_M$ is an embedding, and Lemma 8 finally proves the following:

Lemma 17. $\widehat{\Lambda}_{\mathcal{C}}(M) \subseteq \mathbb{S}^{n-1}$ *is an embedded submanifold of dimension* $dim\widehat{\Lambda}_{\mathcal{C}}(M) = rank_q\Lambda_{\mathcal{C}}$.

If choosing the construction before instead of the compact ball $B_\varepsilon(\varphi(q))$, the open ball $\overset{\circ}{B}_\varepsilon(\varphi(q))$, it is $U := \varphi^{-1}\left(\overset{\circ}{B}_\varepsilon(\varphi(q))\right)$ open in M, i.e., $U \in \mathfrak{U}_q$ is an neighborhood of q open in $\widehat{\mathbb{P}}$, i.e., $q \in U \subseteq_{open} M$. $\widehat{\Lambda}_{\mathcal{C}}|_M$ is an embedding, especially a homeomorphism onto the image, i.e., $\widehat{\Lambda}_{\mathcal{C}}(U) \subseteq_{open} \widehat{\Lambda}_{\mathcal{C}}(M)$. The lemma before showed that $\widehat{\Lambda}_{\mathcal{C}}(M) \subseteq \mathbb{S}^{n-1}$ is an embedded submanifold.

Moreover, an open set $O \subseteq_{\text{open}} N$ of a manifold N is again a manifold with $dimO = dimN$ (Note 5b). The inclusion $\iota : O \hookrightarrow N$ is the identity in proper charts (U, φ) of N and corresponding charts $(O, \varphi|_O)$ of O (w.l.o.g $O \subseteq U$), and thus, it is an immersion. With $O \subseteq_{\text{open}} N$, the inclusion is also a homeomorphism onto $O = \iota(O)$. In summary, $\iota : O \hookrightarrow N$ is an embedding.

With $\widehat{\Lambda}_{\mathcal{C}}(U) \subseteq_{open} \widehat{\Lambda}_{\mathcal{C}}(M)$, the latter proves that $\widehat{\Lambda}_{\mathcal{C}}(U)$ is an embedded submanifold of the embedded submanifold $\widehat{\Lambda}_{\mathcal{C}}(M) \subseteq \mathbb{S}^{n-1}$ (Lemma 8). As a consequence, $\widehat{\Lambda}_{\mathcal{C}}(U)$ is an embedded submanifold of $\mathbb{S}^{n-1}$(Note 8e). Furthermore, $dim\widehat{\Lambda}_{\mathcal{C}}(U) = dim\widehat{\Lambda}_{\mathcal{C}}(M) = rank_q\widehat{\Lambda}_{\mathcal{C}}$. Since the overall construction is valid for each $p \in \widehat{\mathbb{P}}$ the following has been proven:

Note 16. *For each* $p \in \widehat{\mathbb{P}}$ *there exists an open neighborhood of* p, *i.e.,*$p \in U \subseteq_{open} \widehat{\mathbb{P}}$, *such that* $\widehat{\Lambda}_{\mathcal{C}}|_U$ *is an embedding, and* $\widehat{\Lambda}_{\mathcal{C}}(U) \subseteq \mathbb{S}^{n-1}$ *is an embedded submanifold with* $dim\widehat{\Lambda}_{\mathcal{C}}(U) = rank_q\widehat{\Lambda}_{\mathcal{C}}$.

Both, U and $\widehat{\Lambda}_{\mathcal{C}}(U)$ are differentiable manifolds without boundary of the same dimension. Moreover, by construction $\widehat{\Lambda}_{\mathcal{C}}|_U : U \to \widehat{\Lambda}_{\mathcal{C}}(U)$ is differentiable with bijective differential $d\left(\widehat{\Lambda}_{\mathcal{C}}|_U\right)_p$. The Inverse Function Theorem (Theorem 2) then implies the following:

Note 17. *For each $p \in \widehat{\mathbb{P}}$ there exists an open neighborhood U_p of p, i.e.,$p \in U_p \subseteq_{open} \widehat{\mathbb{P}}$, such that $\widehat{\Lambda}_{\mathcal{C}}|_{U_p} : U_p \to \widehat{\Lambda}_{\mathcal{C}}(U_p)$ is a diffeomorphism.*

5.3. An Attempt to Extend Locally Embedded Circuit Manifolds

Many local topological properties, i.e., properties valid in a neighborhood of a point, can be extended to connected components. For example, a function that is locally constant, is constant on each connected component. Central for the corresponding proofs is Lemma 2. In the case of locally constant functions, its use is as follows:

Let X be a connected topological space, and let $f : X \to \mathbb{R}$ be locally constant, i.e., each point $x \in X$ has an open neighborhood $x \in U_x \subseteq_{open} X$ and a number $c_x \in \mathbb{R}$ such that $f|_{U_x} = c_x$. Furthermore, if two such neighborhoods U_x, U_y intersect, the corresponding numbers c_x, c_y are equal: for $z \in U_x \cap U_y$ it is $f(z) = c_x$ and $f(z) = c_y$, i.e., $c_x = c_y$ and, thus, $f|_{U_x} = f|_{U_y}$. Let $\mathcal{U} = \{U_x \mid x \in X\}$ be an open cover consisting of such neighborhoods, and choose two arbitrary points $a, b \in X$. According to Lemma 2, there is an open chain $\{U_1, \ldots, U_n\} \subseteq \mathcal{U}$ connecting a and b. Because $U_i \cap U_{i+1} \neq \varnothing$ it is $f|_{U_i} = f|_{U_{i+1}}$, which implies that $f|_{U_i} = f|_{U_j}$ for any $1 \leq i, j \leq n$. Thus, $f|_{U_1} = f|_{U_n}$, i.e., $f(a) = f(b)$. Because a, b are arbitrary points from X, f is constant.

In our context, the question at hand is whether the rank of a differentiable map (which can locally only increase according to Lemma 4) can globally only increase on a connected component. In order to mimic the proof before, we make the following assumption on the parameter space.

Assumption 2. *For each $q \in \mathbb{P}$ and each $1 \leq r \leq k$ let $\overline{\mathbb{P}}_{q,r} := \mathbb{R}^r_q \cap \mathbb{P}$ be connected., i.e., especially $q \in \overline{\mathbb{P}}_{q,r}$.*

More precisely, it suffices that $\overline{\mathbb{P}}_{q,r}$ has a finite number of connected components, and for the construction that follows, the connected component containing q is chosen. Thus, w.l.o.g., we can assume that $\overline{\mathbb{P}}_{q,r}$ is connected. This is directly the case in many practical situations, e.g., if $\mathbb{P}$ is an appropriate (see the discussion at the end of Section 4.2) semi-open hypercube (see Figure 21). Similarly, a k-dimensional torus is an example of a parameter space $\mathbb{P}$ where $\overline{\mathbb{P}}_{q,r}$ consists of more than one but a finite number of connected components.

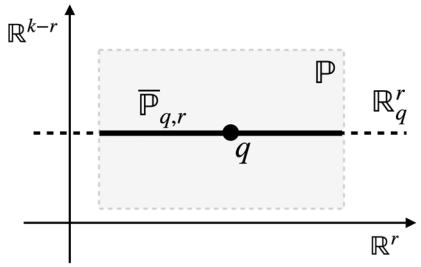


Figure 21. Slicing the Parameter Space.

Let $q \in \overline{\mathbb{P}}_{q,r_q}$ with $r_q := rank_q \Lambda_{\mathcal{C}}$. According to the construction of Section 5.1, there exists an open neighborhood of q with $q \in U_q \subseteq_{open} \overline{\mathbb{P}}_{q,r_q}$ such that $\Lambda_{\mathcal{C}}|_{U_q}$ is an immersion (Lemma 15), i.e., $\Lambda_{\mathcal{C}}|_{U_q}$ has constant rank r_q in U_q.

For $x, y \in \overline{\mathbb{P}}_{q,r_q}$ choose $U_x, U_y \subseteq \mathbb{P}$ as before, such that $\Lambda_{\mathcal{C}}|_{U_x}$ has constant rank r_x in U_x and that $\Lambda_{\mathcal{C}}|_{U_y}$ has constant rank r_y in U_y. Note that $x \in U_x \subseteq_{open} \overline{\mathbb{P}}_{x,r_x}$ but not necessarily $U_x \subseteq_{open} \overline{\mathbb{P}}_{q,r_q}$ or $U_x \subseteq_{open} \mathbb{P}$, similar for U_y. Figure 22 depicts this situation where $\overline{\mathbb{P}}_{q,r_q}$ is two-dimensional, U_x one-dimensional, and U_y three-dimensional. Thus, it cannot be guaranteed that an open cover of $\overline{\mathbb{P}}_{q,r_q}$ exists for which the rank of $\Lambda_{\mathcal{C}}|_U$ is constant for a member U of the open cover. Consequently, even if a $\overline{\mathbb{P}}_{q,r}$ is connected, it

cannot be guaranteed via an argument involving Lemma 2 that the rank is constant on all of $\overline{\mathbb{P}}_{q,r}$.

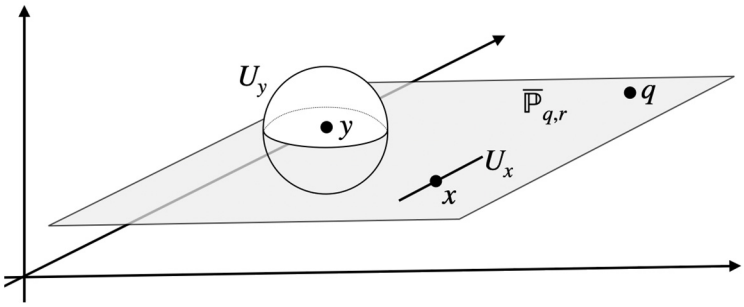


Figure 22. Neighborhoods of Constant Rank of $\Lambda_{\mathcal{C}}$.

There is even a counterexample:

Let $f : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x^2, y^2)$. Then $df_{(x,y)} = \begin{pmatrix} 2x & 0 \\ 0 & 2y \end{pmatrix}$. The rank of $df_{(x,y)}$ is as follows:

- Let $(x, y) = (0, 0)$. Then $df_{(0,0)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow rank_{(0,0)} f = 0$
- Let $y = 0 \wedge x \neq 0$. Then $df_{(x,0)} = \begin{pmatrix} 2x & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow rank_{(x,0)} f = 1$
- Let $y \neq 0 \wedge x = 0$. Then $df_{(0,y)} = \begin{pmatrix} 0 & 0 \\ 0 & 2y \end{pmatrix} \Rightarrow rank_{(0,y)} f = 1$
- Let $y \neq 0 \wedge x \neq 0$. Then $df_{(x,y)} = \begin{pmatrix} 2x & 0 \\ 0 & 2y \end{pmatrix} \Rightarrow rank_{(x,y)} f = 1$

The following figure shows the landscape of the ranks of this map f.

f has rank 0 at the origin (red dot), and rank 1 at the x-axis and the y-axis (the green and blue lines without the origin). Each point of the rest of the plane (gray shaded area) has rank 2. For the point q shown, it is $rank_q f = 2 = r_q$. With a properly chosen parameter space $\mathbb{P}$ (e.g., the dark gray shaded rectangle), the set $\overline{\mathbb{P}}_{q,2}$ is as depicted. $\overline{\mathbb{P}}_{q,2}$ is a manifold without boundary, $\overline{\mathbb{P}}_{q,2}$ is connected, and $dim\overline{\mathbb{P}}_{q,2} = 2$; i.e., $\overline{\mathbb{P}}_{q,2}$ satisfies Assumptions 1 and 2. However, the rank of f is not constant on $\overline{\mathbb{P}}_{q,2}$.

5.4. Determining "Large" Circuit Manifolds

The counterexample before (and Figure 23) indicates that within a set $\overline{\mathbb{P}}_{q,r_q}$ "large" areas A exist in which the rank $rank\Lambda_{\mathcal{C}}|_A = r_q$ of $\Lambda_{\mathcal{C}}|_A$ is constant, i.e., $\Lambda_{\mathcal{C}}|_A$ is an immersion. For example, with $\mathcal{U} = \left\{ U_x \subseteq_{open} \overline{\mathbb{P}}_{q,r_q} \mid rank\Lambda_{\mathcal{C}}|_{U_x} = r_q \right\} \neq \varnothing$ define $A := \bigcup_{U \in \mathcal{U}} U$ and A is then such a "large" area.

Locally, $\Lambda_{\mathcal{C}}|_A$ is also injective (see the proof of Lemma 16). Thus, $\Lambda_{\mathcal{C}}(A)$ is locally an immersed submanifold of $\mathbb{S}^{n-1}$. The domain of injectivity can be extended in concrete situations by analyzing $\Lambda_{\mathcal{C}}(A)$. Here is an example of such an analysis:

Example: *Let $\mathbb{P} = \mathbb{R}$ and let $\mathcal{C}(x) = e^{ixH}$ for a hermitian matrix H. A set $A \subseteq \mathbb{R}$ has to be determined such that $\mathcal{C}(x)$ is injective. Since every hermitian matrix is normal by definition ([34] Definition 18.1) and each normal matrix is diagonalizable ([34] Theorem 18.2), it is $H = Tdiag\left(e^{i\lambda_1}, \cdots, e^{i\lambda_n}\right)T^*$ for a unitary matrix U and the real eigenvalues $\lambda_1, \ldots, \lambda_n$. Consequently, it is $e^{ixH} = Tdiag\left(e^{ix\lambda_1}, \cdots, e^{ix\lambda_n}\right)T^*$. Finding the set of parameters for which $\mathcal{C}(x) = e^{ixH}$ is injective means to determine when $e^{ixH} = e^{iyH}$ implies $x = y$. Thus, we need to find $X \subseteq \mathbb{R}$ such that $e^{ixH} = e^{iyH} \Leftrightarrow x = y$ for $x, y \in X$.*

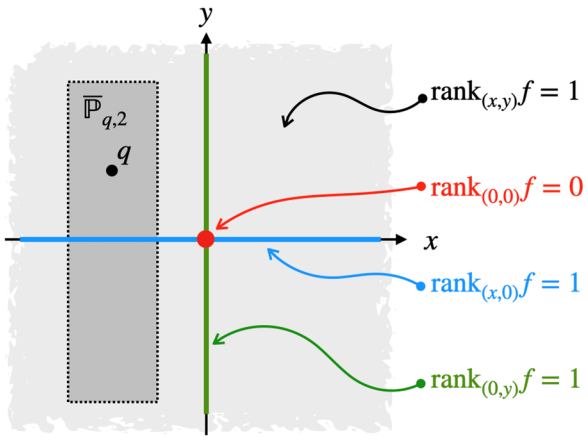


Figure 23. Landscape of Ranks of $f(x,y) = (x^2, y^2)$.

$diag(e^{ix\lambda_1}, \ldots, e^{ix\lambda_n})$, and thus e^{ixH}, may be periodic—which implies non-injectivity: If $\lambda_1, \ldots \lambda_n \in \mathbb{Q}$, then there exists a $T > 0$ such that $e^{iHx} = e^{iH(x+T)}$ for each x.

Assume $x \neq y$ with $e^{ixH} = e^{iyH}$. Then, $e^{i(x-y)H} = I$. Thus, the spectrum of $(x-y)H$ must be in $2\pi\mathbb{Z}$, i.e., $\lambda_j(x-y) \in 2\pi\mathbb{Z}$ for each eigenvalue λ_j of H. If the spectrum of H fulfills

$$\forall T > 0 : \sigma(H) = \{\lambda_1, \ldots, \lambda_n\} \not\subseteq \frac{2\pi}{T}\mathbb{Z} \tag{15}$$

then $e^{i(x-y)H} \neq I$, i.e., $e^{ixH} \neq e^{iyH}$: $\mathcal{C}(x)$ is injective. Thus, if H fulfills condition (15), $\mathcal{C}$ and, thus, $\Lambda_{\mathcal{C}}$ is injective on all of $\mathbb{R} = \mathbb{P}$.

This example can be abstracted to a high-level procedure of how to determine "large" circuit manifolds. First, the matrix $\Lambda_{\mathcal{C}}$ has to be computed based on the ansatz $\mathcal{C}$; this matrix is used in step (2). For an ansatz like in Equation (5), often used in practice, this is straightforward; otherwise, the steps that proved Lemma 1 in Section 3.1 must be followed. Thenm

1. Fix a point $q \in \mathbb{P}$
2. Determine $r_q := rank_q \Lambda_{\mathcal{C}}$
3. Determine the slice $\overline{\mathbb{P}}_{q,r_q}$
4. Determine $\mathcal{U} = \left\{ U \subseteq_{open} \overline{\mathbb{P}}_{q,r_q} \mid rank \Lambda_{\mathcal{C}}|_U = r_q \right\}$
5. Build $A = \bigcup_{U \in \mathcal{U}} U \subseteq_{open} \overline{\mathbb{P}}_{q,r_q}$
6. Determine the maximal subset $G \subseteq A$ such that $\Lambda_{\mathcal{C}}|_G$ is injective
7. Determine a maximal subset $K \subseteq_{compact} G$.
8. $\Lambda_{\mathcal{C}}(K) \subseteq \mathbb{S}^{n-1}$ is an embedded submanifold

Step (4) and (5) determine a maximal open set $A \subseteq_{open} \overline{\mathbb{P}}_{q,r_q}$ such that $\Lambda_{\mathcal{C}}|_A$ has constant rank $r_q = dim \overline{\mathbb{P}}_{q,r_q}$, i.e., $\Lambda_{\mathcal{C}}|_A$ is an immersion. Since $\Lambda_{\mathcal{C}}|_A$ is also locally injective (proof of Lemma 16), step (6) succeeds but is highly dependent on the ansatz $\mathcal{C}$ (see the example before). After step (6) we know that $\Lambda_{\mathcal{C}}|_G$ is an injective immersion, i.e., $\Lambda_{\mathcal{C}}(G)$ is an immersed submanifold of $\mathbb{S}^{n-1}$ (Definition 13). And after step (7) we know that $\Lambda_{\mathcal{C}}(K) \subseteq \mathbb{S}^{n-1}$ is an embedded submanifold (Lemma 6, Lemma 8).

In [53], an open-source package called QMetric is introduced that computes several metrics (especially expressivity) of a parameterized quantum circuit. A description of how expressivity, according to the unitary approach, is measured is given in [31].

6. Conclusions and Future Work

The literature about the expressivity of parameterized quantum circuits requires a lot of background in differential topology, which makes it hard to comprehend. This contribution provides the corresponding background in a single place with detailed references

to textbooks and seminal papers, which may be consulted for a much deeper dive into the domain. Similarly, statements about properties of parameterized quantum circuits are often not proved in the literature, or a proof is only indicated. This contribution provides proof of statements about dimensional expressivity (to be precise: proof for local versions of such properties). Also, counterexamples are given, and limits are pointed out (e.g., by highlighting the importance of singularities).

Section 5 clearly reveals the need for future work: Conditions under which local embeddings can be extended to (global) embedding in the case of parameterized quantum circuits are needed, especially if the circuits have the form of Equation (5), which is practically relevant. Furthermore, the applicability in practical situations of the procedure sketched in Section 5.4 must be evaluated in practice.

Author Contributions: Writing—original draft, F.L.; Writing—review and editing, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

At the end of Section 4.5 it is shown that $df_p(v) = Df(p)(v) = D_v f(p)$, i.e., the differential applied to v is the directional derivative of f in the direction of v. By definition, the directional derivative is $D_v f(p) := lim_{t\to 0}\frac{f(p+tv)-f(p)}{t}$, i.e., $df_p(v) = D_v f(p) = lim_{t\to 0}\frac{f(p+tv)-f(p)}{t}$.

Thus, for $f(X) := X^*X$ we obtain the following:

$$\begin{aligned} df_A(V) &= \lim_{t\to 0}\tfrac{f(A+tV)-f(A)}{t} \\ &= \lim_{t\to 0}\tfrac{(A+tV)^*(A+tV)-A^*A}{t} \\ &= \lim_{t\to 0}\tfrac{(A^*+tV^*)(A+tV)-A^*A}{t} \\ &= \lim_{t\to 0}\tfrac{A^*A+tA^*V+tV^*A+t^2V^*V-A^*A}{t} \\ &= \lim_{t\to 0}\tfrac{t(A^*V+V^*A+tV^*V)}{t} \\ &= \lim_{t\to 0}(A^*V + V^*A + tV^*V) \\ &= A^*V + V^*A \end{aligned}$$

This proves the claim $df_A(V) = A^*V + V^*A$.

Appendix B

We sketch a proof that $f :]-\pi, +\pi[\to \mathbb{R}^2,\ x \mapsto (sin2x, sinx)$ is injective. Assume $s, t \in]-\pi, +\pi[$ and $f(s) = f(t)$. Thus:

$$\begin{aligned} & sin2s = sin2t \wedge sins = sint \\ \overset{(A)}{\Leftrightarrow}\ & 2sinscoss = 2sintcost \wedge sins = sint \\ \overset{(B)}{\Leftrightarrow}\ & sinscoss = sintcost \wedge sins = sint \\ \overset{(C)}{\Leftrightarrow}\ & coss = cost \wedge sins = sint \end{aligned}$$

(A) is because of $sin2x = 2sinxcosx$, (B) is the division by "2", and (C) is because of $sins = sint$ and division by $sins$, w.l.o.g. $sins \neq 0$ (otherwise $s = t = 0$ because of $s, t \in]-\pi, +\pi[$, which implies injectivity).

Next, injectivity of both, $sin :]-\pi/2, +\pi/2[\to \mathbb{R}$ and $cos :]0, +\pi[\to \mathbb{R}$ is considered, thus: $coss = cost$ for $t \in]0, \pi[$ implies that $s \in]-\pi, 0]$; and vice versa $t \in]-\pi, 0[$ implies that $s \in [0, \pi[$. Otherwise $sins = sint$ for $t \in]-\pi/2, +\pi/2[$ implies $s \geq \pi/2$ or $s \leq -\pi/2$.

Now, a case distinction is performed. Case 1: $t \in]0, \pi/2[$ implies both, $s \in]-\pi, 0]$ as well as $s \leq -\pi/2$. But for $t \in]0, \pi/2[$ and $s \in]-\pi, -\pi/2[$ it is $sint \neq sins$. The other cases follow analogously.

References

1. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2016.
2. Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S.C.; Endo, S.; Fujii, K.; McClean, J.R.; Mitarai, K.; Yuan, X.; Cincio, L.; et al. Variational Quantum Algorithms. *Nat. Rev. Phys.* **2021**, *3*, 625–644. [CrossRef]
3. Preskill, J. Quantum Computing in the NISQ era and beyond. *Quantum* **2018**, *2*, 79. [CrossRef]
4. Huang, H.-Y.; Broughton, M.; Mohseni, M.; Babbush, R.; Boixo, S.; Neven, H.; McClean, J.R. Power of data in quantum machine learning. *Nat. Commun.* **2021**, *12*, 2631. [CrossRef]
5. Beer, K.; Bondarenko, D.; Farrelly, T.; Osborne, T.J.; Salzmann, R.; Wolf, R. Efficient Learning for Deep Quantum Neural Networks. *Nat. Commun.* **2020**, *11*, 808. [CrossRef]
6. Sim, S.; Johnson, P.D.; Aspuru-Guzik, A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Adv. Quantum Technol.* **2019**, *2*, 1900070. [CrossRef]
7. Funke, L.; Hartung, T.; Jansen, K.; Kühn, S.; Stornati, P. Dimensional Expressivity Analysis of Parametric Quantum Circuits. *Quantum* **2021**, *5*, 422. [CrossRef]
8. McClean, J.R.; Romero, J.; Babbush, R.; Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **2016**, *18*, 023023. [CrossRef]
9. Blekos, K.; Brand, D.; Ceschini, A.; Chou, C.-H.; Li, R.-H.; Pandya, K.; Summer, A. A Review on Quantum Approximate Optimization Algorithm and its Variants. *Phys. Rep.* **2024**, *1068*, 1–66. [CrossRef]
10. Nocedal, J.; Wright, S.J. *Numerical Optimization*; Springer: Berlin/Heidelberg, Germany, 2006.
11. Xu, W.; Zhou, R.-G.; Li, Y.; Zhang, X. Towards an efficient variational quantum algorithm for solving linear equations. *Commun. Theor. Phys.* **2024**, *76*, 115103. [CrossRef]
12. Rodríguez, G.; Alonso, M.; García-Santiago, X.; Gil, G.B. *Energy Distribution Optimization Using Variational Quantum Algorithms*; Galicia Institute of Technology (ITG): A Coruña, Spain, 2024. [CrossRef]
13. Jaschk, D.; Jaksch, D.; Givi, P.; Daley, A.J.; Rung, T. VariationalQuantum Algorithms for Computational Fluid Dynamics. *AIAA J.* **2023**, *61*, 1885–1894.
14. Leymann, F.; Barzen, J. *The Bitter Truth About Gate-Based Quantum Algorithms in the NISQ Era*; Quantum Science and Technology; IOP Publishing Ltd.: Bristol, UK, 2020.
15. Weder, B.; Barzen, J.; Leymann, F.; Zimmermann, M. Hybrid Quantum Applications Need Two Orchestrations in Superposition: A Software Architecture Perspective. In Proceedings of the 18th IEEE International Conference on Web Services, Madrid, Spain, 20–24 April 2009; Association for Computing Machinery: New York, NY, USA, 2021.
16. Truger, F.; Barzen, J.; Leymann, F.; Obst, J. Warm-Starting the VQE with Approximate Complex Amplitude Encoding. In Proceedings of the 1st International Conference on Quantum Software (IQSOFT 2025), Bilbao, Spain, 10–12 June 2025.
17. Macaluso, A.; Clissa, L.; Lodi, S.; Sartori, C. A Variational Algorithm for Quantum Neural Networks. *Comput. Sci.-ICCS 2020* **2020**, *12142*, 591–604.
18. Morales, M.E.S.; Biamonte, J.D.; Zimborás, Z. On the Universality of the Quantum Approximate Optimization Algorithm. *Quantum Inf. Process.* **2020**, *19*, 291. [CrossRef]
19. Holmes, Z.; Sharma, K.; Cerezo, M.; Coles, P.J. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum* **2022**, *3*, 010313. [CrossRef]
20. Bengtsson, I.; Zyczkowski, K. *Geometry of Quantum States*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2017.
21. Gil-Fuster, E.; Eisert, J.; Dunjko, V. On the expressivity of embedding quantum kernels. *Mach. Learn. Sci. Technol.* **2024**, *5*, 025003. [CrossRef]
22. Gorodski, C. *Smooth Manifolds*; Springer: Berlin/Heidelberg, Germany, 2020.
23. Lee, J.M. *Introduction to Smooth Manifolds*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2013.
24. Du, Y.; Tu, Z.; Yuan, X.; Tao, D. Efficient Measure for the Expressivity of Variational Quantum Algorithms. *Phys. Rev. Lett.* **2022**, *128*, 080506. [CrossRef]
25. Ghosh, T.; Mandal, A.; Banerjee, S.; Mukherjee, N.; Panigrahi, P.K. Lower bound of the expressibility of ansatzes for Variational Quantum Algorithms. *arXiv* **2025**, arXiv:2311.01330v2.

26. Yao, J. Learning to Maximize Quantum Neural Network Expressivity via Effective Rank. *arXiv* **2025**, arXiv:2506.15375v2. [CrossRef]
27. Liu, J.; Yuan, H.; Lu, X.-M.; Wang, X. Quantum Fisher information matrix and multiparameter estimation. *J. Phys. A Math. Theor* **2020**, *53*, 023001. [CrossRef]
28. Yu, Z.; Chen, Q.; Jiao, Y.; Li, Y.; Lu, X.; Wang, X.; Yang, J.Z. Non-asymptotic Approximation Error Bounds of Parameterized Quantum Circuits. *arXiv* **2024**, arXiv:2310.07528v2.
29. Mhiri, H.; Monbroussou, L.; Herrero-Gonzalez, M.; Thabet, S.; Kashefi, E.; Landman, J. Constrained and Vanishing Expressivity of Quantum Fourier Models. *arXiv* **2024**, arXiv:2403.09417v1. [CrossRef]
30. Azado, P.C.; Correr, G.I.; Drinko, A.; Medina, I.; Canabarro, A.; Soares-Pinto, D.O. Expressibility, Entangling Power and Quantum Average Causal Effect for Causally Indefinite Circuits. *Phys. Rev. A* **2025**, *111*, 042620. [CrossRef]
31. Liu, Y.; Kaneko, K.; Baba, K.; Koyama, J.; Kimura, K.; Takeda, N. Analysis of Parameterized Quantum Circuits: On the Connection Between Expressibility and Types of Quantum Gates. *arXiv* **2024**, arXiv:2408.01036v2. [CrossRef]
32. Correr, G.I.; Medina, I.; Azado, P.C.; Drinko, A.; Soares-Pinto, D.O. Characterizing Randomness in Parameterized Quantum Circuits Through Expressibility. *arXiv* **2025**, arXiv:2405.02265v2.
33. Ragone, M.; Bakalov, B.N.; Sauvage, F.; Kemper, A.F.; Marrero, C.O.; Larocca, M.; Cerezo, M. A Lie Algebraic Theory of Barren Plateaus for Deep Parameterized Quantum Circuits. *Nat. Commun.* **2024**, *15*, 7172. [CrossRef]
34. Liesen, J.; Mehrmann, V. *Linear Algebra*; Springer: Berlin/Heidelberg, Germany, 2015.
35. Parthasarathy, K. *Topology*; Springer: Berlin/Heidelberg, Germany, 2022.
36. von Querenburg, B. *Mengentheoretische Topologie*; Springer: Berlin/Heidelberg, Germany, 2001. (In German)
37. Hirsch, M.W. *Differential Topology*; Springer: Berlin/Heidelberg, Germany, 1976.
38. Lopez, R. *Point-Set Topology*; Springer: Berlin/Heidelberg, Germany, 2024.
39. Lee, J.M. *Introduction to Riemannian Manifolds*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2018; p. S21.
40. Lueck, W.; Macko, T. *Surgery Theory*; Springer: Berlin/Heidelberg, Germany, 2024.
41. Arnold, V.I. *Catastrophe Theory*; Springer: Berlin/Heidelberg, Germany, 1984.
42. Bröcker, T.; Lander, L. *Differentiable Germs and Catastrophes*; Cambridge University Press: Cambridge, UK, 1975.
43. Joyce, D. A generalization of manifolds with corners. *Adv. Math.* **2016**, *299*, 760–862. [CrossRef]
44. Whitney, H. Differentiable Manifolds. *Ann. Math.* **1936**, *37*, 3. [CrossRef]
45. Adhikari, A.; Adhikari, M.R. *Basic Topology 2*; Springer: Berlin/Heidelberg, Germany, 2022.
46. Riemann, B. Ueber die Hypothesen, welche der Geometrie zu Grunde liegen. In *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen*; Dieterichsche Buchhandlung: Mainz, Germany, 1868; Volume 13. Available online: https://www.emis.de/classics/Riemann/Geom.pdf (accessed on 31 August 2025). English Translation: On the Hypotheses Which Lie at the Bases of Geometry. Available online: https://www.maths.tcd.ie/pub/HistMath/People/Riemann/Geom/WKCGeom.html (accessed on 31 August 2025).
47. Tu, L.W. *An Introduction to Manifolds*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2011.
48. San Martin, L.A.B. *Lie Groups*; Springer: Berlin/Heidelberg, Germany, 2021.
49. Koch, R. *Compact Lie Groups*; University of Oregon: Eugene, OR, USA, 2022; p. S21.
50. Axler, S. *Measure, Integration & Real Analysis*; Springer: Berlin/Heidelberg, Germany, 2020.
51. Grigoryan, A. *Analysis on Manifolds*; University of Bielefeld: Bielefeld, Germany, 2024.
52. Alexandrino, M.M.; Bettiol, R.G. Introduction to Lie groups, isometric and adjoint actions and some generalizations. *arXiv* **2010**, arXiv:0901.2374v3. [CrossRef]
53. Illésová, S.; Rybotycki, T.; Beseda, M. QMetric: Benchmarking Quantum Neural Networks Across Circuits, Features, and Training Dimensions. *arXiv* **2025**, arXiv:2506.23765v2. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.