MAPPING EQUIVALENCE UNDER ITERATIVE DYNAMICS FOR SYMBOLIC SEQUENCES

Liming Wang

Dept. of Electrical Engineering
Columbia University
liming@ee.columbia.edu

Dan Schonfeld

Dept. of Electrical & Computer Engineering
University of Illinois at Chicago
dans@uic.edu

ABSTRACT

In order to employ powerful tools in signal processing to analyze symbolic sequences, a mapping is commonly applied first to transform the symbolic sequences to numerical sequences. Therefore it is important to investigate the role of the mapping in the final analysis results. The concepts of mapping equivalence and related theory have been proposed previously for the case where the data is processed by an operator for only once. However, many operators such as de-noising filter, smoothing filter and certain algorithm may be utilized multiple times. In this paper, we extend the concepts of mapping equivalence to the case of iterations of operators. We provide various theoretical results on determining the equivalence of two mappings. We also establish the connection of analysis robustness to the proposed mapping equivalence concepts. We provide numerical examples to illustrate and justify our theoretical results.

Index Terms— Mapping equivalence, complex dynamics, Fatou set, Julia set.

1. INTRODUCTION

Signal processing is a common and powerful strategy for exploring the information contained in various signal sequences. Most tools in signal processing are proposed based on the fact that the signal takes values in a field $\mathbb{F}$ or a vector space over $\mathbb{F}$. Typical choice of the field is $\mathbb{R}$ or $\mathbb{C}$. Although some tools such as Fourier analysis can be defined on group [1], for most cases, signal processing tools can not be easily and elegantly extended to accommodate other algebraic structures. Meanwhile, many kinds of signals we encounter are in symbolic form which may not have any meaningful algebraic structure. For example, in genomic signal processing, it is difficult to interpret the amino acids or DNA symbols as an algebraic structure on which classical signal processing can be implemented directly. In order to cope with such problem, we can perform a mapping for the symbolic sequence to get a numerical sequence. Analysis conclusion of the original symbolic sequence is drawn based on the transformed numerical sequence. Such approach is popular in genomic signal processing, pattern analysis, biostatistics, etc. Since distinct choices of mappings may lead to contradictory analysis results, it is important to recognize the role of the mapping to the final result. In previous work, the authors [2, 3] explored the relationship between the mapping and the final results by proposing concepts of mapping equivalence. We showed the equivalent mapping for certain classes of operators including correlation function and Fourier transform. Our result suggests that two analysis results can only be compared if two mappings used are equivalent.

The mapping equivalence and consistency concepts proposed in [2] mainly focus on the case that an operator is applied only once. Typical examples in this case include Fourier transform, correlation function, etc. However, it is common that the signal may undergo various pre- or post-processing to eliminate noise or emphasize certain features of interests [4]. In these cases, operators as de-noising filter, smoothing filter and certain algorithm may iterate multiple times for the data. The equivalence issue rises naturally in these situations. Another interesting and important interpretation of the equivalence issue which we will elaborate later is the robustness of the mapping chosen. Specifically, we are interested to find whether small error in acquisition of data leads to small error in the final results.

In this paper, we extend the concepts of mapping equivalence and consistency to the case of iterations of operator. We derive several theoretical properties for the consistency based on complex dynamics. In Section 2, we first define various concepts of equivalence and introduce the concepts of Fatou and Julia Set. We establish the connection between consistency to Fatou and Julia set. In Section 3, we establish the connection of mapping consistency to the robustness. In Section 4, we present experimental results which illustrate the theoretical results.

2. MAPPING EQUIVALENCE UNDER ITERATIONS

In this paper, we consider the case where the symbolic sequence is mapped to complex domain. Let the symbolic sequence be denoted as $\{a_i\}_{i=0}^{N-1}$, where $a_i \in \mathcal{A}$. $\mathcal{A}$ is the collection of symbols. f is a mapping from $\mathcal{A}^N$ to $\mathbb{C}^N$, i.e. $f : \{a_i\}_{i=0}^{N-1} \mapsto z,\ z \in \mathbb{C}^N$. Note that we allow the mapping to be time-variant. Namely, the same symbol can be mapping to different numbers at different time. Therefore, for a given symbolic sequence and a mapping f, the mapped numerical sequence corresponds to one point in $\mathbb{C}^N$, which is denoted as z_f. Under many circumstances, the filtering or certain algorithm can be written explicitly as a complex function $\mathbf{\Phi}$ [5]. Thus the multiple applications of the filtering or algorithm can be formulated as the iterative composition $\mathbf{\Phi}^{\circ n}(z_f)$, where $\mathbf{\Phi}^{\circ n}(z_f)$ denotes the n-th function compositions. More specifically, let $\mathbf{\Phi} : \mathbb{C}^N \to \mathbb{C}^N$ be a holomorphic (analytic) operator. We will assume $\mathbf{\Phi}$ is a polynomial, i.e. $(\mathbf{\Phi}(z))_i = P_i(z_1, z_2, ..., z_N), i = 1, ..., N$, where $(\cdot)_i$ denotes the i-th entry and P_i is a polynomial. We note that this assumption is not very restrictive, since by Taylor's theorem, any holomorphic map can be approximated by polynomials.

Given a symbolic sequence and two mappings f and g, after we apply the operator iteratively for each numerical sequence z_f and z_g, we say that the two analysis results are consistent if they show certain form of similarity. In that case, we say that the two mappings are equivalent. The reasoning is that for one symbolic sequence whose information is already fixed, any discrepancy in the analysis results should be caused by different choices of mappings. Therefore, it is not reasonable to compare the inconsistent results unless the mappings are equivalent.

We first define different concepts of mapping equivalence.

Definition 1. *Given a symbol sequence and two mapping f and g, f and g are asymptotically equivalent if*

$$\lim_{n\to\infty} \|\mathbf{\Phi}^{\circ n}(z_f) - \mathbf{\Phi}^{\circ n}(z_g)\| = 0. \tag{1}$$

f and g are called M-boundedly equivalent, if

$$\sup_n \|\mathbf{\Phi}^{\circ n}(z_f) - \mathbf{\Phi}^{\circ n}(z_g)\| < M. \tag{2}$$

f and g are called n-th equivalent, if

$$\|\mathbf{\Phi}^{\circ n}(z_f) - \mathbf{\Phi}^{\circ n}(z_g)\| = 0. \tag{3}$$

Now the mapping equivalence under iterative dynamics can be investigated as two problems. First, given two mappings, determine whether they are equivalent and the type of mapping equivalence. Second, given one mapping, we need to determine its equivalent mapping class, i.e., all the mappings which are equivalent to the given one.

In study of iteration dynamics, the concepts of Fatou and Julia set play a fundamental role. There are several different definitions of Fatou and Julia set [6, 7]. We will use the definition below. Before that, we first introduce the notion of *normality*.

Definition 2 ([8]). *A collection of holomorphic map $\mathcal{F}$ is called normal if every infinite sequence of maps from $\mathcal{F}$ either has a locally uniformly convergent subsequence or a subsequence diverges locally uniformly.*

Definition 3 ([6]). *The domain of normality F of $\mathcal{F} = \{\mathbf{\Phi}^{\circ n}\}$ is called Fatou set. Its complement*

$$J = \mathbb{C}^N \backslash F \tag{4}$$

is called Julia set.

We define the basin of infinity as set of all points which have norms go to infinity under iteration.

The connected components of Julia (Fatou) set are called Julia (Fatou) components.

We will see later the Julia set represents the chaotic behaved points and points in Fatou set show rational behavior. We can show the following propositions about the Fatou set. The proof of the following two propositions in one-dimensional case appears in [6].

Proposition 1. *A point z is in Fatou set if z is in the basin of infinity.*

Proposition 2. *Fatou (Julia) component is invariant. i.e. the operator maps one component onto another component.*

For z_f and z_g, if only one of them is in the basin of infinity, it is obviously that f and g are not boundedly or asymptotically equivalent. If both are in the basin of infinity, although theoretically we can examine the equivalence, however, from computational point of view, the points diverge very fast under polynomial iterations. After a few rounds of iterations, the numerical results will overflow. In this case, the equivalence or even analysis result turns out to be meaningless. Braverman and Yampolsky [9] showed the Julia set of certain types of polynomial can not be computed by any Oracle Turing Machine. If one of the point is in Julia set, it may not computationally capable to figure out the equivalent mapping class of the given maps. We classify all the mappings falling in all above situations as the *computationally chaotic mapping class*. In general, from computational point of view, it is futile or meaningless to find the equivalent mappings of the element in this class. Therefore, the only interesting case left would be if both points are in Fatou set.

As for the simplest case where $\mathbf{\Phi}$ is affine, we can show the following results for equivalent mapping.

Theorem 1. *If $\mathbf{\Phi}(z) = Az + b$, all mappings are asymptotically equivalent for any symbolic sequence if and only if the spectral radius $\rho(A) < 1$.*

All mappings are boundedly equivalent for any symbolic sequence if and only if either $\rho(A) < 1$ or $\rho(A) = 1$ and all the eigenvalues have index ≤ 1.

In order to evaluate more complicate case, we need to introduce a new metric other than Eculidean metric, which simplifies the analysis. For complex manifold, one can construct a pseudo metric called *Kobayashi pseudo metric*. A complex manifold is called *hyperbolic* if the Kobayashi pseudo metric

d_K is a metric. [10] is referred for the details of construction and properties of Kobayashi metric and hyperbolic manifold.

We can show the following theorems on certain class of operator.

Theorem 2. *If $\mathbf{\Phi}$ is non-degenerate homogenous polynomial, i.e. $\mathbf{\Phi}$ is homogenous and $\mathbf{\Phi}^{-1}(0) = 0$, then its Fatou component is hyperbolic. Moreover, for any two mappings z_f and z_g in a Fatou component, we have*

$$d_K(\mathbf{\Phi}(z_f), \mathbf{\Phi}(z_g)) \leq d_K(z_f, z_g). \tag{5}$$

In other words, any two mappings are boundly equivalent in Kobayashi metric. Meanwhile, any two mappings in a Fatou component are also boundly equivalent in Euclidean metric.

The following theorem can be established for the asymptotical equivalence.

Theorem 3. *If $\mathbf{\Phi}$ is non-degenerate homogenous polynomial, U is a Fatou component, and $\mathbf{\Phi}(U) = U$, if $d_K(\mathbf{\Phi}(x), \mathbf{\Phi}(y)) < d_K(x, y)$ for any distinct $x, y \in U$, then there exists a unique fixed point in U and any z_f and z_g in U are asymptotically equivalent.*

3. ROBUSTNESS OF MAPPING UNDER ITERATIONS

In this section, we establish the connection of robustness of a mapping to the mapping equivalence. In analysis of the symbolic sequences, the noise and error will inevitably affect the final analysis result. Specifically, the issue whether the analysis result is robust to perturbation of the signal is vital for the credibility of the result. At the same time, we can always view the perturbation on the symbolic sequence as a perturbation of the mapping we choose. Therefore, the robustness of the analysis is equivalent to the robustness of the mapping. Namely, whether small perturbation of the mapping will lead to small perturbation of the final analysis result. We first give a rigorous definition of robustness of mapping.

Definition 4. *Given a symbolic sequence, a mapping f is called robust, if for any $\delta > 0$, there exists $\epsilon > 0$ such that for any point z_g in the ball of radius ϵ, centered at z_f we have*

$$\|\mathbf{\Phi}^{\circ n}(z_f) - \mathbf{\Phi}^{\circ n}(z_g)\| < \delta, \forall n \in \mathbb{N} \tag{6}$$

In other words, all the mappings in the ball are δ-boundedly equivalent.

We can show the following results about the robustness of mapping.

Theorem 4. *If f is not in basin of infinity, then a mapping f is robust if and only if z_f is in Fatou set.*

From theorem 4, we can see that Fatou set represents the good-behaved mappings. Any mapping close enough will be a boundedly equivalent mapping. On the contrary, for the mapping in the Julia set, no matter how close the mapping is, it may not even be a boundedly equivalent mapping.

4. EXPERIMENTAL RESULTS

We choose the symbolic sequences as two DNA sequences AD169 and rhodopsin gene sequence. We consider the operator $\mathbf{\Phi}$ as a non-linear smoothing filter defined as follow,

$$\mathbf{\Phi}(z_1, z_2, ..., z_N) = (\frac{{z_1}^2 + {z_2}^2}{2}, ..., \frac{{z_i}^2 + {z_{i+1}}^2}{2}, ..., \frac{{z_N}^2}{2}) \tag{7}$$

The symbolic sequence is mapping according to the following map,

$$\tilde{f}(a) = \begin{cases} 1 & \text{if } a = \mathtt{A} \\ -1 & \text{if } a = \mathtt{T} \\ i & \text{if } a = \mathtt{G} \\ -i & \text{if } a = \mathtt{C} \end{cases} \tag{8}$$

As before, we denote the induced mapping point as z_f.

In Fig. 1, we show the slices of Julia and Fatou set of $\mathbf{\Phi}$ at $(z, 1, 1, ..., 1)$, $(z, i, i, ..., i)$ and $(z, 0.25 + 0.75i, ..., 0.25 + 0.75i)$. Julia set commonly possesses a fractal shape and could be connected or disconnected.

It can be shown the Fatou component U containing origin satisfies all the assumptions in theorem 3. Therefore any two points in U are asymptotically equivalent. We consider the following perturbed signal $\tilde{f}'$,

$$\tilde{f}' = \tilde{f} + \Delta z \tag{9}$$

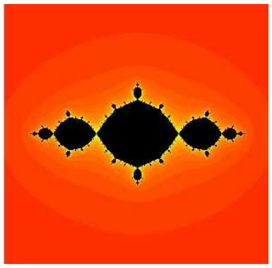
where $\Delta z \in \mathbb{C}$. In Fig. 2, we show the slice of the Fatou component U with z_f at origin and Δz is varying in the ball of radius 0.1, centered at 0. The white area is in the Fatou component U.

In Fig. 3 (a), we show how the Euclidean distance for two arbitrarily chosen mappings which are in the same Fatou component U changes with the number of iterations for the AD169 sequence. As we can see that the distance converges to 0 with the increase of number of iterations as predicted. In Fig. 3 (b), we show how the Euclidean distance for two non-equivalent mappings changes with the number of iterations. One is in previous Fatou component U and the other is very close to previous one but not in U. As we can see that the distance diverges with the increase of number of iterations.

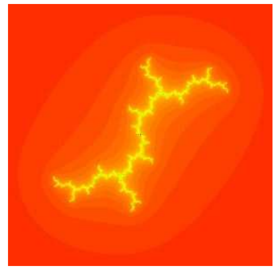
5. REFERENCES

[1] G. Caire, R. L. Grossman, and H. V. Poor, "Wavelet transforms associated with finite cyclic groups," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1157–1166, 1993.

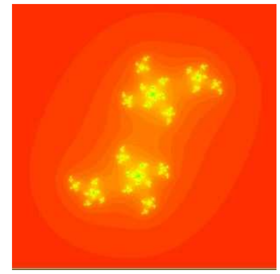
[2] L. Wang and D. Schonfeld, "Mapping equivalence for symbolic sequences: Theory and applications," *IEEE Trans. on Signal Processing,*, vol. 57, no. 12, pp. 4895 –4905, Dec. 2009.



(a) Julia Set and Fatou Set

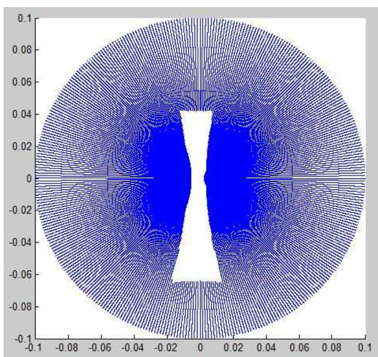


(b) Julia Set and Fatou Set

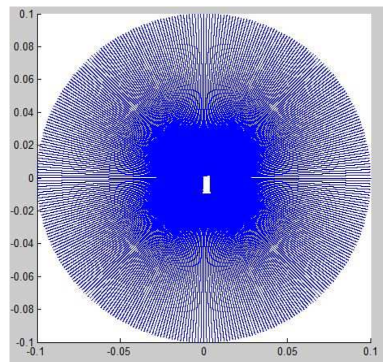


(c) Julia Set and Fatou Set

Fig. 1. Slices of the Fatou and Julia Set of at $(z, 1, 1, ..., 1)$, $(z, i, i, ..., i)$ and $(z, 0.25 + 0.75i, ..., 0.25 + 0.75i)$ respectively. The Julia set is represented as the golden color. The red and black color represent the Fatou Set.

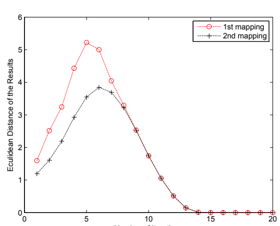


(a) Illustration of slice of the Fatou component U for AD169 sequence.

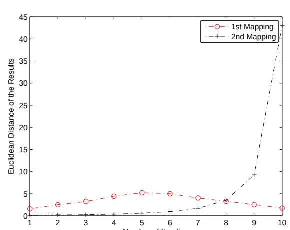


(b) Illustration of slice of the Fatou component U for rhodopsin sequence.

Fig. 2. Asymptotical equivalence: (a), (b) show the slice of Fatou component U containing z_f for AD169 sequence and rhodopsin gene sequence respectively. The origin represents z_f and the central white area is in U.



(a)



(b)

Fig. 3. Illustrations of how Euclidean distances for two mappings change with the number of iterations: (a) The case that two mappings are in the same Fatou component U. (b) The case that one mapping is in the Fatou component U and the other is not.

[3] L. Wang and D. Schonfeld, "Consistency in representation and transformation of genomic sequences," in *IEEE International Workshop on Genomic Signal Processing and Statistics,*. IEEE, 2009, pp. 1–4.

[4] W. Wang and D. H. Johnson, "Computing linear transforms of symbolic signals," *IEEE Trans. on Signal Processing*, vol. 50, no. 3, pp. 628–635, March 2002.

[5] F. Carravetta, A. Germani, and M. Raimondi, "Polynomial filtering for linear discrete time non-Gaussian systems," *SIAM Journal on Control and Optimization*, vol. 34, pp. 1666, 1996.

[6] J. Milnor, *Dynamics in One Complex Variable*, vol. 160 of *Annals of Mathematics Studies*, Princeton University Press, 3rd edition, 2006.

[7] L. Carleson and T. W. Gamelin, *Complex Dynamics*, Springer-Verlag, 1993.

[8] L. Ahlfors, *Complex Analysis*, McGraw-Hill, 3rd edition, 1979.

[9] M. Braverman and M. Yampolsky, "Constructing non-computable Julia sets," in *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*. ACM, 2007, pp. 709–716.

[10] S. Kobayashi, *Hyperbolic Manifolds and Holomorphic Mappings*, Marcel Dekker, New York, 1970.