



Audio Engineering Society
Convention Paper 10667

Presented at the 155th Convention
2023 October 25–27, New York, USA

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (<http://www.aes.org/e-lib>), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Detection of phase alignment and polarity in drum tracks

Alek Weidman¹, John Sweeney¹, and Wesley A. Bulla¹

¹*Belmont University, MS Audio Engineering*

Correspondence should be addressed to Alek Weidman (alek.weidman@bruins.belmont.edu)

ABSTRACT

A time-shift applied to individual tracks that removes timing differences between microphones, called “phase alignment,” is frequently promoted as a way to improve the clarity and definition of live-recorded drum tracks. Common techniques include manual and automated micro-timing adjustments and switching the electrical polarity of “problem” tracks. This study aimed to determine if there was a clear audible difference between an original and corrected recording. Using a paired comparison test, listeners were asked whether two audio samples were the same or different, and later, asked for individual preference between the two samples. Evidence here questions the tacit assumption that time-shift techniques have the claimed influence to greatly improve, or even appreciably alter, the observed quality of a drum mix.

1 Introduction

Using more than one microphone to record a single source introduces the possibility of constructive or destructive phase interplay between microphone signals [1]. In drum recordings, the most common sources of potentially destructive phase issues are the different arrival times of drum set components (snare, toms, bass drum, etc.) at individual microphones placed around the drum kit [2]. Given specific stimuli, listeners can reliably detect changes in sound quality caused by phase-oriented interactions [3].

In the popular press, many recording engineers advocate for so-called “phase aligning” of drum tracks as a way to improve the clarity and definition of the recording [4, 5]. Common “corrections” include switching the electrical polarity of individual tracks and timing adjustments between tracks ranging from as low as 10 micro-sec for 96 kHz sample rate to 30 ms (or greater) to account for distances between the drums and overhead or room microphones. The focus of this study was

to determine if listeners could differentiate between drum tracks that had and had not been “phase aligned” and if listeners would recognize and favor a snare drum track with a change in polarity on the bottom-head microphone.

2 Prior Art

Drum kit recording has evolved into an art, craft, and science where techniques range from using a few microphones strategically placed in proximity of the drum kit to utilizing any number of microphones to capture each individual component of the instrument (i.e., bass drum, snare drum, each tom drum, cymbals, etc.). Various best practices and specialized techniques have been documented and detailed over the years in both the popular press, academic journals, and in several well-known instructional texts such as [6] and [7].

In common practice, a microphone is assigned to each individual drum and cymbals are captured as a group

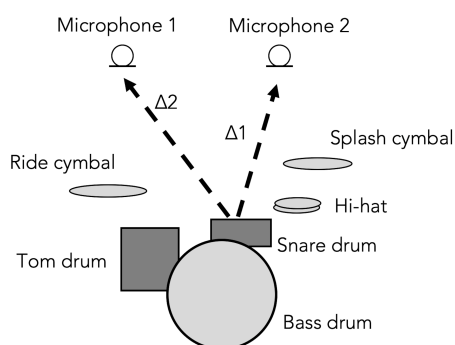


Fig. 1: Illustration of a single signal at two microphones. The difference in distance of each microphone from the source will govern the interaction of the two signals when summed.

via “overhead” microphones placed some distance above the drum kit. Often, two microphones are used to capture the top and bottom of the snare drum and the front and back of the bass drum. Additional microphones may also be employed to capture reverberation from the room in which the drum set is being recorded.

The management of individual source signals at multiple microphones introduces unique challenges. When two microphones that carry the same signal are blended electronically, the resultant signal is the sum of the two. If the two microphones are not precisely equidistant from the source, the sonic quality at the output can be, for better or worse, greatly altered through comb filtering [8] commonly referred to as “phasing.” Fig. 1 illustrates a basic scenario with drum recording where this problem may arise. The outcome of the interaction is distant-dependent on the following formula:

$$c/|\Delta 1 - \Delta 2| = f \quad (1)$$

where: c is the speed of sound, Δ is the distance to each microphone from the source, and f is the lowest accentuated frequency in a set of frequencies that will change the harmonic content of the drum’s sound [9]. Multiples of f will add together and frequencies half-way between the multiples (i.e., f_1, f_2, f_3 , etc.) will subtract. For an excellent visualization of the effect of comb filtering on the resultant blended signal due to changes in Δ , see [10] figures 1-3.

One of the techniques used to address comb filtering is to place microphones such as the overhead microphones 1 and 2 in fig. 1 close together in what is called a coincident relationship where all sources arrive at the two microphones synchronously [11] thus eliminating any time difference between the two signals. Another is to follow the so-called “3-to-1 rule” [12, 13] where any secondary microphones in the vicinity of the source instrument are placed at least three times the distance of the primary microphone from the source. This technique ensures the level of the secondary signal, relative to the primary signal, is low enough that any constructive and destructive effects are minimized and masked by the relative strength of the primary signal.

With a snare, tom, and bass drum, because the primary microphones are usually placed only a few inches from the drum head it is easy to place microphones at distances that take advantage of the 3-to-1 rule. With regard to drum recording, close-mic techniques are effective at minimizing a noticeable interaction between microphones when combined for the final blend of the drum kit. Problems are claimed to arise when overhead microphones are introduced into the blend or when multiple microphones are placed on a single drum. For example, again using fig. 1 as a reference, the sound of the snare drum will arrive at its primary microphone well before it arrives at the overhead microphones and some drums (e.g., snare) are typically recorded with a microphone on each side of the drum or with a microphone placed inside and outside of the drum (e.g., bass).

“Correcting” the signals through time adjustments so that the snare drum, overhead microphone 1 and microphone 2 all start at the same time (i.e., “line up”) is claimed to give the resultant blend “more transient smack” and thus will better “cut through” the mix [14]. In the case of the top and bottom snare microphones which are on opposite sides of the drum, there is the assumption that the batter and snare heads act together as one membrane always moving in the same direction and the recorded signals are always inverted. Consequently, it is common practice to reverse the polarity of one of the microphones so that it will “match” the other and thus give a “fuller” sound [15]. However, it is worth noting this assumption and practice discounts the fact that drum heads have multimodal resonances that result in patterns with considerably more complex relationships [16, 17].

Modern audio engineering, sound recording, and music mixing require that multiple signals are constantly being captured, edited, and re-blended to achieve desired technical and artistic outcomes. Therefore, the effects of summing signals with differing phase and polarity relationships has been well studied and most often applied with careful microphone placement during the recording process [18, 19].

On the other hand, the ability to easily manipulate and compare signal relationship adjustments across the spectrum during the mixing process is a relatively recent innovation and even though recording engineers and technicians have long utilized spatial manipulation of the microphone to make adjustments of the sound during the recording process there is a general lack of empirical confirmation that listeners are highly sensitive to such changes as those currently associated with and recommended for improvement of pre-recorded drum tracks.

The simple listening experiment employed here was designed to establish whether or not listeners could detect a snare drum that had been “phase aligned” by manipulation of the overhead tracks and to determine if listeners would favor a drum track with a change in polarity of the bottom-head microphone track.

3 Methods

The null hypothesis assumed subjects would not be able to detect whether phase alignment or polarity reversal techniques had been applied to a drum set recording. Seven different excerpts, two-bar phrases and a drum fill, were taken from high quality multi-track recordings [20, 21]. Two paired versions of each clip were created where tracks were either phase aligned, remained as originally recorded, or had the bottom track of the snare drum’s electrical polarity reversed.

Utilizing a paired-comparison paradigm [22] subjects were asked to identify whether two tracks played in succession were either the same or different, quantified as correct (1) or incorrect (0). Twelve subjects participated in four randomized and counterbalanced sets comprising forty trials each (twenty same + twenty different).

The first set presented an isolated (ISO) condition using only the top snare and overhead tracks where the overhead tracks were time-aligned (TSA) or not (TSN) with the top snare track. For these samples, only the

Table 1: Sample list with overhead track correction times.

<i>Sample</i>	<i>Artist</i>	<i>Correction(ms)</i>
<i>A</i>	<i>Alchemy</i>	16
<i>B</i>	<i>Cole</i>	37
<i>C</i>	<i>Blaton</i>	10
<i>D</i>	<i>Araujo</i>	9
<i>E</i>	<i>Balazs</i>	8
<i>F</i>	<i>Chalga</i>	12
<i>G</i>	<i>Maciwoda</i>	10

top microphone track was active. Timing adjustments ranged from 8 ms to 37 ms (see table 1 for track list with modification times). The second set incorporated the same modification but with a full blend of all elements of the drum kit (FULL). Sets three, four, and five tested polarity reversal on the bottom snare track in ISO and FULL conditions respectively where the final set presented twenty trials with the bottom snare reversed (BSI) or as recorded (BSO) and asked subjects to identify which clip they preferred.

Subject ages were clustered in the mid-20s. All subjects reported normal hearing, had completed a graduate-level course in critical listening, and some had semi-professional audio engineering training and production experience. Subjects here would generally be considered more experienced than “novice” but not yet “expert” listeners. Subjects were not aware of the purpose of the experiment. Subject responses were recorded and tabulated via Scantron [23] response sheet.

The listening test took place in a group setting in an acoustically treated classroom using high-quality circumaural headphones. Each sample was about five to seven seconds long. Each trial presented two samples with a one second pause in between and a five-second pause allotted for the response time. In total, the experimental trials took about sixty minutes to complete.

4 Results

Data were segregated by whether or not a paired trial repeated a stimulus (SAME) or comprised stimuli with the phase-alignment or polarity treatment applied to one of the stimuli (DIFF). SAME trials served as control for which subject performance could be compared.

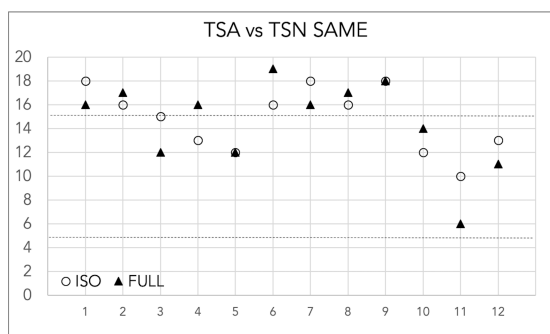


Fig. 2: Plot of response data by subject for the control (SAME) condition where the overhead tracks were time-aligned (TSA) or not (TSN) with the top-snare drum.

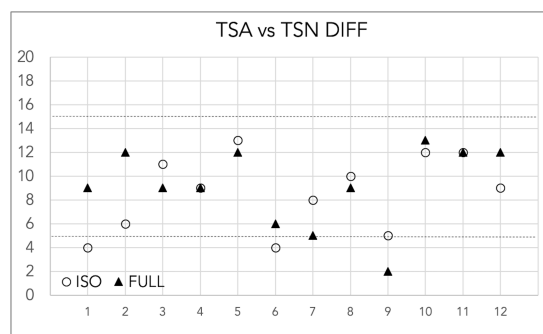


Fig. 4: Plot of response data by subject for the experimental (DIFF) condition where the overhead tracks were time-aligned (TSA) and not time-aligned (TSN) with the top-snare drum.

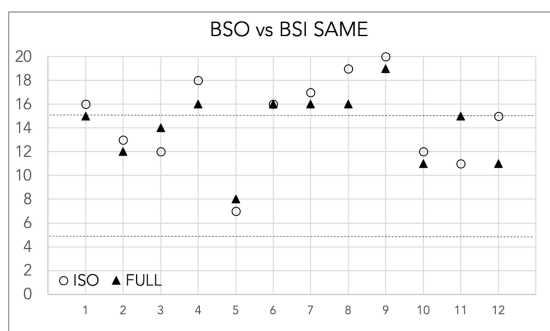


Fig. 3: Plot of response data by subject for the control (SAME) condition where the polarity remained as recorded (BSO) or was reversed (BSI) on the bottom track of the snare drum.

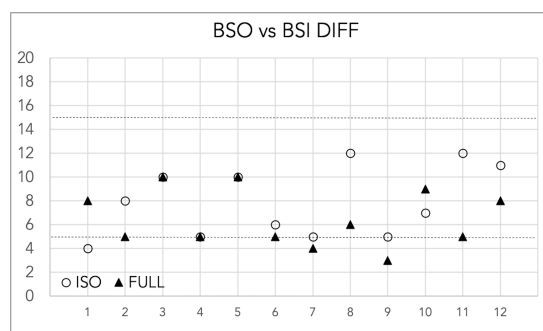


Fig. 5: Plot of response data by subject for the experimental (DIFF) condition where the polarity remained as recorded (BSO) and was reversed (BSI) on the bottom track of the snare drum.

Given the binomial nature of the test design, data were scored separately within each respective comparison group where α criterion $p = .041$ corresponded with seventy-five percent correct or twenty-five percent incorrect responses across twenty trials (i.e., fifteen or five of twenty responses) [24]. In figures 2 through 5, the criterion for reliable detection of a control or experimental variable is shown by the upper and lower dashed lines at fifteen and five correct hits. For the binomial probability [25], these are equivalent outcomes and one must assume successful detection of the treatment variable in either case.

For the SAME condition (Fig. 2 and Fig. 3) all comparison categories surpassed the population criterion of fifty percent or greater [26] where all but two sub-

jects obtained criterion in at least one of the variable categories confirming all but two ($M = 0.83$) subjects (5 and 10) could perform the experimental task.

For the DIFF trials (Fig. 4 and Fig. 5) an even fifty percent of subjects reached criterion in the FULL BSO vs BSI comparison. All other paired comparisons fell short of demonstrating detectability with only three, two, and four subjects respectively able to reach criterion in the TSA vs TSN ISO, TSA vs TSN FULL, and BSO vs BSI ISO trials.

For the BSO vs BSI preference test, results were almost evenly split where 121 ($M = 0.51$) of 240 responses favored the polarity-corrected excerpt and 119 ($M = 0.49$) responses preferred the “as recorded” excerpt.

For the six subjects that reached criterion in the FULL BSO vs BSI test, fifty-six responses ($M = 0.47$) selected a polarity-corrected excerpt and sixty-four responses selected an “as recorded” excerpt ($M = 0.53$). Subject-by-subject analysis did not reveal identifiable patterns suggesting choices here were likely driven by chance.

Analyses also revealed no identifiable patterns for any influence of the seven different excerpts with 33 percent to 52 percent correct discrimination across the five clips used in the phase-alignment trials and 16 percent to 46 percent correct discrimination across the five clips used in the polarity-reverse trials.

5 Discussion

Eight of twelve subjects depicted in Fig. 2 demonstrated an ability to reliably recognize when the two samples comprised the same treatment as either “corrected” with the time-alignment process applied and as originally recorded with no treatment. Six of the eight surpassed the detection criterion in both conditions where the snare and overhead tracks were isolated (ISO) or when the full drum set was playing (FULL). Four subjects failed to reach the criterion in either type of trial.

The plot in Fig. 3 for the top and bottom snare drum polarity trials show a very similar pattern among subject performance. Two subjects, 5 and 10 again failed to reach detection criterion in both the ISO and FULL trials. Additionally, two other subjects (2 and 3) who were successful in the time-align trials also failed to reach detection criterion in the polarity test. Six of twelve subjects demonstrated an ability to reliably recognize the paired stimuli were the same when, in fact, they were the same. In total, again, eight of twelve subjects demonstrated an ability to reliably recognize when the two samples comprised the same treatment in at least one of the playback conditions, ISO or FULL.

Figures 4 and 5 plot data for the experimental trials where one of the samples was the originally recorded track and the other was the altered track, either time-aligned or polarity reversed. When one sample had the overhead tracks time-aligned with the top-snare drum and the other was not, all but four subjects (1, 6, 7, and 9) show guessing performance. As a group, in the time-alignment test, these were the only subjects able to reliably demonstrate an ability to recognize when the two samples were different (Fig. 4). Evidentially,

the task was extremely difficult where only one third of this population demonstrated capability of performing the change.

The outcome for the polarity reversal test (Fig. 5) revealed an interesting result. All seven of the twelve subjects that demonstrated an ability to recognize when the two samples were different in at least one of the playback conditions, executed their responses in reverse order. Regardless, results here clearly show the majority of subjects could recognize when the polarity had been reversed on the bottom snare drum track. The evidence that subjects can detect this change in a stimulus does not carry any surprising element as the resultant effect from the process of blending two signals such as these is simply a drastic change in timber. One would expect most individuals with normal hearing and a basic command of the descriptive language of acoustics to recognize there is something different between these two stimuli.

Results from this experiment suggest that subjects could not hear differences between stimuli in the time-align test but could in the polarity-reverse test. Observations align with previous studies, where discrimination of so called “phase anomalies” has been seen to be highly and often selectively subject-dependent. Here, a single subject (9) demonstrated an ability to correctly detect when the stimuli were either the same or different in all eight experimental conditions presented in this investigation. Preferences for one type of stimuli over the other were split nearly down the middle when looking at the population as a whole. Additionally, when considering subjects that reliably indicated they could hear differences between stimuli, there was no distinct preference for so-called “phase corrected” over the not-corrected samples.

This experiment examined discrimination and preference for time alignment of drum tracks in isolation without additional musical elements. While not tested herein, the authors acknowledge it is possible, one could argue, that differences in individual tracks resulting from the alignment process might alter the perception of the final mix when not in isolation. For example, in an instance where a given track appears to “cut through the mix” possibly when time alignment of certain elements causes a slight increase in the apparent loudness or fullness of the resultant signal. However, because loudness is an obvious and controllable factor it was compensated for within the test design in order

to remove it from potential influence as a confounding factor. It stands to reason that alterations in the apparent quality or character of a signal that would go un-noticed when isolated from the other musical elements of a mix are, more practically speaking, unlikely to also be great enough to cause a given signal to then be noticed when blended into the other elements within a given mix.

The authors also recognize that final mixes are constructed through many small changes that often go unobserved when the listener is unaware of the details behind any specific alteration and that there may be any number of technical reasons for utilizing time-alignment techniques. For example, corrections may be applied because of how phase affects secondary signal processing such as the application of equalization, compression, or limiting as applied to a drum stem or mix bus. This study has considered only those claims associated with correcting the drum track itself and whether those alterations might be readily noticeable by listeners who were, in fact, listening for and expecting a change in the signal.

6 Summary

This study investigated whether or not listeners could reliably determine when a drum track had been phase-corrected or had the polarity reversed on one of the snare drum microphones. Findings from this experiment contradict the widespread belief that time-alignment of drum tracks in a typical studio recording are mandatory and essential fixes in music production.

Results provided no evidence aware listeners could identify when phase-alignment techniques had been applied. Changing the polarity relationship between the top and bottom snare microphones was identified at the minimum population a criterion showing results from a polarity change was noticeable. However, even when listeners could detect a difference in reversal of the polarity on the bottom snare drum track, there was no distinct preference for a so-called “phase corrected” version over the recorded “as is” version.

7 Citations

[1] G. Sigismondi, R. Waller and T. Vear, *Microphone Techniques, Recording*; A Shure Educational Publication, Niles, IL: Shure Inc., (2014).

[2] W. Wadhams, "Chapter 12: Drums," in *Sound Advice: A Musician's Guide to the Recording Studio*, 1st ed., NY, Simon Schuster Macmillan, pp. 135-174, (1990).

[3] M. V. Laitinen, K. Jussila and V. Pulkki, "Effect of phase on the perceived level of bass," in *AES 137th Convention*, Los Angeles, USA, (2014).

[4] D. Thompson, "Drum Overhead Mics Time/Phase Alignment," 13 November 2017. URL: <https://recording.org/forum/microphones/drum-overhead-mics-timephase-alignment>. (Accessed May 2023).

[5] B. Owsinski, *The Recording Engineers Handbook* (4th ed.), Bobby Owsinski Media Group (2017).

[6] W. Wadhams, *Sound Advice: A Musician's Guide to the Recording Studio*, (1st ed.), NY: Simon Schuster Macmillan, (1990).

[7] S. R. Alten, *Audio in Media*, (8th. ed.), CA: Thomson Wadsworth, (2008).

[8] S. R. Alten, "Chapter 8: Signal Processors, Flanging," in *Audio in Media*, (8th. ed.), CA, Thomson Wadsworth, p. 159, (2008).

[9] J. Eargle, "Chapter 5 Microphones, Cancellation in Combining Two or More Microphones," in *fSound Recording* (2nd ed.), New York, NY, Litton Educational Pub, pp. 133-134, (1980).

[10] B. Kruk and A. Sobecki, "Methods of phase-aligning individual instruments recorded with multiple microphones during post-production," in *AES 140th Convention*, Paris, France URL: <https://www.aes.org/e-lib/browse.cfm?elib=18278>, (2016).

[11] J. Eargle, "Stereo Microphone Techniques," in *The Microphone Handbook*, (1st ed.), Plainville, NY, Elar Publishing, pp. 105-133, (1982).

[12] "The Basics About Comb Filtering (and How to Avoid It)," DPA Microphones, Inc, 3 Sept. 2019. URL: <https://www.dpamicrophones.com/mic-university/the-basics-about-comb-filtering-and-how-to-avoid-it>. (Accessed May 2023).

[13] S. R. Alten, "Chapter 9: Studio Production: Talk and Voice-overs, Phase and Polarity," in *Audio in Media*, (8th. ed.), CA, Thomson Wadsworth, pp. 173, 486, (2008).

- [14] H. Robjohns, M. Senior and M. Walker, "Should I be time-aligning my drum tracks?," *Sound-on-Sound Online*, July 2010. [Online]. URL: <https://www.soundonsound.com/sound-advice/q-should-be-time-aligning-my-drum-tracks>. (Accessed May 2023).
- [15] D. Pratt and A. Bourbon, "Life In Between Phase: Understanding and Manipulating Microphone Relationships with Visualisation Tools," in *12th Art of Record Production Conference*, Stockholm, (2019).
- [16] E. K. E. Scott and A. Morrison, "Acoustics of the Modern Jazz Drum Kit," *Acoustics Today*, vol. 18, no. 4, pp. 31-39, (2022).
- [17] T. Rossing, I. Bork, H. Zaho and D. O. Fystrom, "Acoustics of Snare Drums.," *J. Acoust. Soc. of Am.*, vol. 92, <https://doi.org/10.1121/1.404080>, pp. 84-94, (1992).
- [18] D. Rochman, "Common Techniques for Stereo Miking," Shure USA, 23 March 2017. URL: <https://www.shure.com/en-US/performance-production/louder/common-techniques-for-stereo-miking>. (Accessed March 2023).
- [19] DPA, "The Ultimate Guide: Stereo Recording Techniques Setup," DPA Microphones, Longmont, CO 2023. URL: <https://www.dpamicrophones.com/mic-university/stereo-recording-techniques-and-setups-2023>. (Accessed Jan. 2023).
- [20] DPA, "The Ultimate Guide: Stereo Recording Techniques Setup," DPA Microphones, Longmont, CO 2023. URL: <https://www.dpamicrophones.com/mic-university/stereo-recording-techniques-and-setups-2023>. (Accessed Jan. 2023).
- [18] Cambridge Music Technology, "The 'Mixing Secrets' Free Multitrack Download Library," (n.d.). Tracks from: Andrew Cole, (Dead Roses), Ann Blaton, (Waves), Araujo, (The Saga of Harrison Crabfeathers), The Balazs Daniel Boogie Woogie Trio (Own Way To Boogie), Chalga (Szárád A Száj), Pawel Maciwoda (1 Of The Firstz). URL: <https://www.cambridge-mt.com/ms/mtk/>. (Retrieved Oct. 2021).
- [21] Telefunken Elektroakustik, "Studio Sessions," TELEFUNKEN Elektroakustik, South Windsor, CT, Track from Alchemy Microphone Comparison (Multi-Track Session Files, Bedrock LA 2001-2021. URL: <https://www.telefunken-elektroakustik.com/multitracks>. (2001-2021).
- [22] G. V. Michaelson and M. Hardin, "Pairwise Comparisons," in *Encyclopedia of Research Design* (ed. Neil Salkind), Sage Pub. Inc., p. URL: <https://doi.org/10.4135/9781412961288>, (2010).
- [23] Scantron, "Insight 20/30 form no. 223127, GPAS-100Qx5," store.scantron.com, (N.D.).
- [24] "GraphPad, Sign and binomial test," GraphPad Software, Inc, 2023. URL: <https://www.graphpad.com/quickcalcs/binomial2/>. (Accessed May 2023).
- [25] E. W. Weisstein, "Binomial Distribution." *MathWorld – A Wolfram Web Resource*, URL: <https://mathworld.wolfram.com/BinomialDistribution>. (Accessed Sept. 2023).
- [26] J. Blauert, Experimental Procedures, 1.3.1 Psychometric Methods in *Spatial Hearing, The Psychophysics of Human Sound Localization*, London: MIT Press, pp. 14-22 (1997).