PARKER C. FOLSE III (WA Bar No. 24895 – *Pro Hac Vice*)
pfolse@susmangodfrey.com
IAN B. CROSBY (WA Bar No. 28461 – *Pro Hac Vice*)
icrosby@susmangodfrey.com
FLOYD G. SHORT (WA Bar No. 21632 – *Pro Hac Vice*)
fshort@susmangodfrey.com
SUSMAN GODFREY, L.L.P.
1201 Third Avenue, Suite 3800
Seattle, Washington 98101-3000
(206) 516-3880 Tel.
(206) 516-3883 Fax

SPENCER HOSIE (CA Bar No. 101777)
shosie@hosielaw.com
BRUCE WECKER (CA Bar No. 078530)
bwecker@hosielaw.com
HOSIE McARTHUR LLP
One Market, 22nd Floor
San Francisco, CA 94105
(415) 247-6000 Tel.
(415) 247-6001 Fax

Attorneys for Defendant/Counterclaimant
BURST.COM, INC.

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

APPLE COMPUTER, INC., Plaintiff/Counterdefendant, v. BURST.COM, INC., Defendant/Counterclaimant.	CASE NO. C06-00019 MHP

DECLARATION OF ALLEN GERSHO IN SUPPORT OF BURST'S OPPOSITION TO APPLE'S MOTION FOR SUMMARY JUDGMENT OF INVALIDITY BASED ON KRAMER AND KEPLEY PATENTS

I, ALLEN GERSHO, declare:

1. In 1960, I received a B.S. in Electrical Engineering from the Massachusetts Institute of Technology, and I received a Ph.D. in Electrical Engineering from Cornell University in 1963.

2. From 1963 to 1980, I was a member of the technical staff at Bell Laboratories. During that time, I was engaged in fundamental research in signal processing and digital communications. My research included such topics as quantization, delta modulation, ADPCM, digital filtering, and other topics that have direct application to speech and audio compression or digital communications.

3. From 1980 to 1999 I held the position of Professor of Electrical and Computer Engineering at the University of California at Santa Barbara ("UCSB"). During this period, my research focus was compression of speech, audio, image, and video. I supervised and trained many graduate students and visiting scholars and produced twenty Ph.D. graduates, many of whom have since achieved international recognition for contributions to speech, audio, and video compression. I have published over 300 technical papers and have been an editor and reviewer for various engineering journals. From 1984 to 1989, I led a project for NASA to develop new speech compression algorithms for mobile satellite communications.

4. In addition, from 1996-2000, I was CEO of SignalCom, Inc., a consulting company focused primarily in the area of speech coding. SignalCom developed innovative speech compression technologies for Nokia, Qualcomm, and for the National Security Agency. One of our speech coders has since been adopted as a standard for secure voice communication systems by NATO and by the U.S. Department of Defense. From April 2000 through April 2001, I was a digital media architect in the Digital Media Division of Microsoft Corporation. Since January 1999, I have been Professor Emeritus and a Research Professor with the Electrical and Computer Engineering

Department at UCSB. Since 1987, I have been president of VoiceCraft, Inc., a California corporation and have developed and licensed speech and audio coding algorithms. Current licensees include AT&T and Nokia.

5. I hold more than 15 U.S. patents in the areas of speech compression, signal processing, and digital communications. I was elected a Fellow of the IEEE in 1982 and have received numerous awards and professional recognitions for my work in the areas of signal and speech processing and compression. For example, I received the Guillemin-Cauer Prize Paper Award in 1979, the IEEE Centennial Medal in 1984, the IEEE Donald McClellan Meritorious Service Award in 1983, the Circuits and Systems Video Technology Prize Paper Award in 1992, the Ericsson-Nokia Prize Paper Award in 1999, the IEEE Millennium Medal in 2000, and the Society Award of the IEEE Signal Processing Society, 2003. In April 2007 at a ceremony in Honolulu, Hawaii, I was presented with the IEEE James L. Flanagan Speech and Audio Processing Award, 2007.

6. I have co-edited (with B.S. Atal and V. Cuperman) two books on speech and audio coding and authored or co-authored numerous articles and book chapters in the area of speech coding and digital signal compression. I co-authored (with R.M. Gray) the book "Vector Quantization and Signal Compression," published in 1992, an advanced text which has since become an international standard in the area of signal quantization. A list of my publications and curriculum vitae are attached as Exhibit A.

7. I am currently providing expert consulting services for Burst, Inc. ("Burst") in the above mentioned lawsuit. I am being compensated at my standard billing rate of $750 per hour plus expenses. My compensation is not contingent on the testimony that I intend to offer in this case.

8. I have read and studied Burst's four patents, Apple's summary judgment motion, the Kramer and Kepley patents that Apple has asserted as prior art, and the Court's claim construction ruling in this case.

9. In her claim construction expert report for Apple v Burst, dated October 20, 2006, Dr. Sheila Hemami gave an opinion regarding the level of ordinary skill in the art as follows:

> Based upon my knowledge of the development of digital communications and networking and my personal involvement in the development of compression technology, a person of ordinary skill in the art at the time that the patent application leading to the '995 patent was filed would have had an understanding of: (1) digital communication technologies and their available bandwidths, and (2) audio and/or video compression techniques.
>
> In general, a person of ordinary skill in the art would work in the area of digital communication of audio/video source information. A person in this area could be specialized in digital communications having a familiarity with compression technology, or such a person could be specialized in compression technology having a familiarity with digital communications.
>
> Such a person of ordinary skill in the art would have had at least a bachelor's degree in electrical engineering with at least two to three years of experience working on digital communication of audio/video source information.
>
> Alternatively, such a person of ordinary skill in the art would have had a master's degree in electrical engineering with one year of experience working on digital communication of audio/video source information. As another alternative, such a person of ordinary skill in the art would have had a Ph.D. degree in electrical engineering in the area of digital communication of audio/video information.

I agree with Dr. Hemami's opinion.

10. Kramer does not disclose a "transceiver" apparatus having "input means" and "compression means" in a common housing with "random access storage means" and "output means" as required in Claim 1 of the '995 patent. Apple has identified the function of "encod[ing] . . . into digital form" identified at Kramer 3:10 as indicating that "compression means" are present in the Kramer memory card. (By the term "memory card" I am referring to the "portable data processing

and storage card" shown in Figure 1 of the patent.) Kramer clearly states that this encoding occurs "outside the illustrated system." It is clear that this encoding occurs outside of the memory card. One with ordinary skill in the art would understand this statement to mean that the encoding does not occur in either the memory card or the playback unit depicted in Figure 2 of the Kramer patent. Moreover, neither the memory card nor the playback unit is depicted as containing any structure that could perform the referenced encoding. In addition, because the "input means" in '995 Claim 1 receives information prior to compression, based on the sequence in the claim, any corresponding structure in Kramer would necessarily be "outside the illustrated system" as well.

11. Kramer does not employ "compression" as the Court has construed that term to mean "reducing the number of bits." One with ordinary skill in the art would understand the reference to "encod[ing] . . . into digital form" using "DPCM" at Kramer 3:10 to refer to converting an analog signal directly into a digital signal having a DPCM format. Analog implementations of DPCM encoding were well known in the art and widely available at the time of the Kramer patent. See for example the references Keiser85, Bellamy91 and Jayant84 attached as Exhibits B, C and D, respectively, to this declaration. Each of these references shows a DPCM block diagram that specifies the input as analog. Such a conversion does not result in a reduction in the number of bits of a signal because analog information has no bits to reduce.

12. Kramer does not transmit audio faster than real time from card to card as Apple claims. To explain why Kramer does not teach transmission in a burst time period, it is necessary to understand how Kramer uses sub-band coding. In the digital encoding (encoding from analog to digital) of the audio source with sub-band coding, the original analog audio signal is applied to a bank of filters; each filter passes a segment of the acoustic frequency spectrum producing a sub-band audio signal. Each analog sub-band signal is digitally encoded with DPCM to produce a separate bit

stream for that frequency band. The sub-band bit streams are then multiplexed, to interleave the bits into one composite serial bit stream. Control information is added and the resulting data stream is what is supplied to the memory of the memory card. Although Kramer does not explicitly state this, one with ordinary skill in the art would understand that the patent illustrates the use of sub-band coding for the example of 100 sub-bands. The data signal coming from memory contains interleaved data for each of 100 sub-bands. Hence, (without sub-sampling) this data specifies 100 sub-band audio signals, each a filtered version of the original audio signal and each at real time. Hence the bit rate from the memory must be at least 100 times the rate needed to reproduce the sound of any one of these sub-band audio signals. The mixer combines these sub-band audio signals into one composite real time audio signal.

13. Considering the part of the memory card consisting of the multiplexer, decoders, and analog signal mixer as one black box, this box has as its input the serial data stream from the memory and as its output the real time reproduced audio signal. This black box itself has no memory; hence the speed at which the data comes from the memory into the black box is exactly at the rate needed to reproduce at the output of the box the final audio work in real time. If the stored data represented a 3.5 minute song, it would take 3.5 minutes for this data to pass out of the memory into the black box in order to produce 3.5 minutes of analog audio output. Therefore, one of ordinary skill would understand that the speed at which the data is delivered from the memory is the speed needed for real time reproduction of the sound and does not represent a burst mode of data flow. Hence for direct digital output of the memory card, this same bit stream is being transmitted to another memory card or to the replay unit. In either case, it is a real-time transmission and not a burst transmission.

14. This is further substantiated by Kramer's statements:

> In another mode of use, the system can be programmed so that the output is of the data in its digital form; the controller 34 then instruct

> the output register 30 to pass the data from the memory 22 directly to the output multiplexer 32, without passing through the recoders [*sic* decoders] 26; such a digital output can be used, as mentioned above, as input into another memory, e.g. of another portable card of the invention. (5:5-12)

This statement says that the digital data can be transmitted out of the memory card without passing through the decoders and can be fed directly to another memory card. The description of the replay unit shown in Fig. 2 of the patent shows that when data is sent out of the memory card in digital form to a replay unit, it can be replayed directly rather than necessitate storage or buffering (as would be needed in the replay unit if data arrived in a burst mode). Furthermore, the use of the word "directly" in the above statement would be understood by one with ordinary skill in the art to imply that the same data (with the same data rate) that was applied to the aforementioned black box can be rerouted to bypass the black box and proceed directly to the output multiplexer for transmission in digital form without undergoing any rate changing operation. Hence, the data in digital form is transmitted at real-time in order to replay the audio in real time. If data were transmitted at a speed, for example, 10 times faster than real time, then all the data representing a 5 minute song would arrive in 30 seconds at the replay unit. In order to play this song the replay unit would necessarily need to store this data (e.g. 5 x 60 seconds at say 300,000 bits per second is 90 million bits) in some sizable memory unit on the replay unit. But the replay unit has no memory, only an input register whose sole described function is to "remove control data from the stream of data". It has never been suggested that this input register can operate as a memory for data storage as does the bubble memory on the memory card. Such a register is likely to hold a very small number of bits. Nowhere is there any indication that the memory on the memory card has the capability of reading out data at a multiplicity of data transmission speeds depending on where this data is destined. There is no indication in the patent to support the idea that the card to card transfer rate is different or faster than the rate for feeding data to the demultiplexer.

15. The "bubble memory" as configured in Kramer is expressly disclosed as a serial-access storage device. Kramer explicitly contrasts this type of memory with other types of memory that allow immediate recall of any given segment of information. One of ordinary skill in the art would understand these statements to expressly distinguish the memory discussed in Kramer from "random-access storage" as that term is understood by a person of skill in the art and has been construed by the Court. Moreover, one of ordinary skill in the art would not consider the memory in Kramer to be "coupled" to "compression means," inasmuch as Kramer performs only analog coding, which is not compression. Furthermore this encoding does not occur in any part of the card in which the memory in Kramer is situated.

16. Kramer does not expressly teach receiving any information from a computer, nor is such teaching inherent. The referenced "computer information" could be received by the Kramer system for storage in the memory card from some other storage medium such as a computer disk and not necessarily directly from a computer. Examples of devices besides computers in use around the time of Kramer from which the memory card could have received computer information include an IBM 711 punched card reader, an IBM 726 magnetic tape reader, or a bar code reader with digital output in RS232 serial data format. Bar codes with the current standard were first used in 1972. Since this data is ultimately destined for a computer, the data is computer information. The storage device can provide mobile transfer of the data to a remote computer. People taking inventory of products in a supermarket could conceivably use a bar code reader with a Kramer storage device for holding the data. These examples of computer information are digital in origin.

17. Kramer also does not disclose receiving computer generated audio/video source information. Kramer expressly says that the referenced computer information is "other than sound." For example, scientific data such as analog seismic data taken from a seismic sensor could be digitized and then be considered "computer information" since it is destined for a computer (rather than received from a computer) that will subsequently analyze this data. Kramer does not disclose or process video information. Consequently, one of ordinary skill in the art would conclude that the "computer information" reference in Kramer (6:31-33) is merely computer data with no audio or video component and that this data is not necessarily received from a computer.

18. I understand the "removable storage medium" limitations of the Burst patents to refer to a storage medium that can be inserted and removed from the claimed Burst transceiver. The only storage medium disclosed in Kramer is the memory card of Figure 1. Kramer does not disclose inserting or removing the memory card of Figure 1 in or from the system that does the encoding from analog to digital. Although Kramer indicates that the player unit depicted in Figure 2 can have a slot for receipt of a card, the player unit is not a transceiver as claimed in the Burst patents. Apple claims that the memory card itself constitutes a transceiver, but Kramer never discloses or suggests inserting or removing a memory card into or from another memory card. Thus one with ordinary skill in the art would not understand Kramer to teach inserting or removing a removable storage medium into or from a transceiver.

19. Kepley does not expressly or inherently disclose the audio compression methods required by the Court's construction of "compression means." Digital encoding and compression of speech (voice) has been studied extensively over the past forty years. See for example, my review paper on advances in speech and audio coding, Gersho94 attached as Exhibit E. A rich variety of approaches and algorithms have been developed particularly in the 1970s and 1980s, prior to the

filing of the AT&T Kepley patent. Most of the speech coding algorithms were developed during this period at AT&T Bell Laboratories (where I worked from 1963 to 1980).

The Kepley patent refers to the use of speech coding, saying

> bandwidth compression (compress the voice data from 64k bits/s down to 16k bits/s) and expansion; silence compression (encode the length of long silences so that the encoded length value rather than the actual silent interval can be stored on disk) and expansion

Here bandwidth compression refers to what is usually called speech coding while silence compression is an accessory technique.

20. The patent does not teach what particular algorithm should actually be performed in the voice processor for performing bandwidth compression or silence compression. In fact, there is a rich variety of speech coding algorithms that were known prior to the filing of the Kepley patent. Below, I identify a few such algorithms that do not fit into the category of "compression means." Specifically, none of these algorithms reduce the bit rate by comparing two or more samples and encoding certain differences between those samples. Each of these algorithms can be applied to PCM encoded speech at 64 kb/s to reduce the bit rate to 16 kb/s or lower. These algorithms are based on the use of one or both of two key approaches: *Vector Quantization* (VQ) and *Analysis by Synthesis* (AbS).

21. In *waveform VQ* (also known as *Vector PCM*), the voice signal is partitioned into segments or blocks. See: Gray84, AbutGray 1982, Tao1984, DavidsonGersho86a attached as Exhibits F, G, H and I, respectively. Each block of consecutive samples (called a "vector") is compared with a predetermined codebook of stored vectors and based on a measure of disparity or distortion, the best matching vector in the stored codebook is determined and the address or index of the selected vector is the digital code sent to the decoder. The decoder has a copy of the codebook

and is able to regenerate an approximation to the original speech waveform by extracting the indicated vector for each block. Although a comparison between a signal vector and a codebook vector is performed, no differences of any kind are being coded.

22. *Hierarchical Vector Quantization* (HVQ) [Gersho84 attached as Exhibit J] is an enhanced version of Vector PCM. It allows much longer blocks of samples to be encoded by partitioning the blocks into sub-blocks and extracting a measure of the energy of each block. The vector of energies is coded with VQ and the quantized energy levels are used to normalize the sub-blocks. The normalized sub-blocks are then individually encoded with VQ. HVQ does not do comparisons between samples and no differences of any kind are coded.

23. Several different algorithms for speech coding belong to the family of AbyS speech coding. They each reproduce (or " synthesize") speech from the digital representation in such a way that the speech sounds to the human listener very close to the original while the reproduced waveform does not necessarily appear similar to the original speech waveform. In each such algorithm, the encoder contains a replica of the decoder. The encoder takes a block of input samples and performs a trial-and-error search process to determine which of a large set of binary code words will do the best job in reproducing that input block. The search is based on a perceptual measure that compares the input block with the possible synthesized blocks. Although the comparison may involve forming differences between samples, these differences are not coded. They serve only to aid the search process to find the best input codeword.

24. In *trellis coding* of speech waveforms, [Stewart82 attached as Exhibit K], the decoder may consist of a shift register combined with a nonlinear operation on the received data bits in the register, or a table-lookup producing a binary word that is applied to a digital to analog converter. The decoded speech samples are formed from the outputs of this converter. At the encoder, one or

more input speech samples is used to perform a trellis search to find a group of bits that cause the decoder to produce the best match to the input samples. Comparisons are made between the input samples and the various candidate samples that might be produced by the decoder, but no differences are coded.

25. In *Multipulse LPC Coding* (Singhal84 attached as Exhibit L) and *Regular Pulse Excitation Coding* (Kroon86 attached as Exhibit M) the decoder synthesizes speech by applying a binary word (a group of bits) to an excitation generator that computes a block of signal amplitudes, called an excitation. This excitation is then applied to a synthesis filter whose output is a block of synthesized speech samples. The excitation is constrained to have a few isolated nonzero amplitudes called pulses while the remaining sample amplitudes are zero. No differences are coded.

26. *Code Excited Linear Prediction* (CELP) (Schroeder85 attached as Exhibit N) and *Vector Excitation Coding* (VXC) (DavidsonGersho86b attached as Exhibit O) are two different names for the same speech coding technique where the excitation is generated by a VQ codebook containing a large number of predetermined candidate excitation vectors. The decoder receives a binary word that supplies the address in the codebook to produce the chosen excitation vector. No differences are coded.

27. Silence compression is based on identifying time intervals where the speech is absent and operates by measuring the comparative energy level of segments of speech. It does not compare individual samples of speech and it does not encode speech samples nor does it encode any differences between speech samples. Silence compression merely identifies silent intervals so that the associated speech encoder can avoid wasting bits encoding silent segments. The only comparisons performed by the silence compression algorithm are to compare energy levels of different segments of speech. Furthermore, I agree with the conclusion of Dr. Hemami that the use of

silence compression disclosed in Kepley would not result in a substantial reduction in bit rate. In fact, silence compression has much greater value in full duplex voice conversations (i.e., two speakers talking to each other) where one speaker is on average silent for half the time. In that case a significant reduction in bit rate is achievable.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct, and that, based on my knowledge, qualifications, and experience, I am prepared to testify on the subjects addressed herein.

Allen Gersho June 7, 2007

Allen Gersho

References (listed in the order cited in the declaration)

Exhibit B
[Keiser85] Digital Telephony and Network Integration
Keiser, B.E. and Strange, E., Van Nostrand Reinhold, 1985; p.41, fig. 4-3

Exhibit C
[Bellamy91] Digital Telephony, 2nd edition
Bellamy, J. John Wiley & Sons, 1991; p. 130, fig. 3.27.

Exhibit D
[Jayant84] Digital Coding of Waveforms; Principles and Applications to Speech and Video
Jayant, N.S., Noll, P.
Prentice-Hall, Inc. 1984; p. 256, Fig. 6.2(a).

Exhibit E
[Gersho94] Advances in speech and audio coding
Gersho, A. Proc. IEEE,
Volume 82, no. 6, June 1994.

Exhibit F
[Gray84] Vector quantization
Gray, R.; ASSP Magazine, IEEE [see also IEEE Signal Processing Magazine]
Volume 1, Issue 2, Part 1, Apr 1984 Page(s): 4 – 29.

Exhibit G
[AbutGray 1982] Vector quantization of speech and speech-like waveforms
Abut, H.; Gray, R.; Rebolledo, G.;
Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on
Volume 30, Issue 3, Jun 1982 Page(s):423 – 435

Exhibit H
[Tao1984], Hardware Realization of Waveform Vector Quantizers
Tao, B.; Abut, H.; Gray, R.;
Selected Areas in Communications, IEEE Journal on
Volume 2, Issue 2, Mar 1984 Page(s):343 – 352

Exhibit I
[DavidsonGersho86a] Application of a VLSI Vector Quantization Processor to Real-Time Speech Coding
Davidson, G.; Gersho, A.;

Selected Areas in Communications, IEEE Journal on
Volume 4, Issue 1, Jan 1986 Page(s):112 - 124
Volume 10, Apr 1985 Page(s):1437 – 1440

Exhibit J
[Gersho84] Hierarchical vector quantization of speech with dynamic codebook allocation, Gersho, A.; Shoham, Y.;

Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.
Volume 9, Part 1, Mar 1984 Page(s):416 - 419

Exhibit K
[Stewart82] The Design of Trellis Waveform Coders
Stewart, L.; Gray, R.; Linde, Y.;
Communications, IEEE Transactions on [legacy, pre - 1988]
Volume 30, Issue 4, Apr 1982 Page(s):702 - 710

Exhibit L
[Singhal84] Improving performance of multi-pulse LPC coders at low bit rates
Singhal, S.; Atal, B.;
Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.
Volume 9, Part 1, Mar 1984 Page(s):9 - 12

Exhibit M
[Kroon86] Regular-pulse excitation--A novel approach to effective and efficient multipulse coding of speech
Kroon, P.; Deprettere, E.; Sluyter, R.;
Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on
Volume 34, Issue 5, Oct 1986 Page(s):1054 – 1063

Exhibit N
[Schroeder85] Code-excited linear prediction(CELP): High-quality speech at very low bit rates
Schroeder, M.; Atal, B.;
Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '85.
Volume 10, Apr 1985 Page(s):937 – 940

Exhibit O
[DavidsonGersho86b] Complexity reduction methods for vector excitation coding

Davidson, G.; Gersho, A.;
Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86.
Volume 11, Apr 1986 Page(s):3055 – 3058