PARKER C. FOLSE III (WA Bar No. 24895 – *Admitted Pro Hac Vice*)
pfolse@susmangodfrey.com
IAN B. CROSBY (WA Bar No. 28461 – *Admitted Pro Hac Vice*)
icrosby@susmangodfrey.com
FLOYD G. SHORT (WA Bar No. 21632 – *Admitted 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

(*additional attorneys listed on signature page)*

Attorneys for Defendant
BURST.COM, INC.

UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA (SAN FRANCISCO)

APPLE COMPUTER, INC.,	§	Case No. 3:06-CV-00019 MHP
	§	
Plaintiff/Counterdefendant,	§	**DECLARATION OF ALLEN GERSHO**
	§	**IN OPPOSITION TO APPLE'S**
	§	**SECOND MOTION FOR**
	§	**SUMMARY JUDGMENT OF**
BURST.COM, INC.,	§	**INVALIDITY**
	§	
Defendant/Counterclaimant.	§	
	§	Hon. Marilyn Hall Patel
	§	
	§	Complaint: January 4, 2006
	§	Trial: February 26, 2008

I, Allen Gersho, submit this Declaration pursuant to 28 U.S.C. § 1746 and declare as follows:

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 AT&T 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 1998 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 largely on 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 compression algorithms. Current licensees include AT&T, Lucent, Agere, Avaya, 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.

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 the book "Vector Quantization and Signal Compression," published in 1992, an advanced text which has been translated into Japanese and has become an internationally recognized benchmark reference in the area of signal quantization and compression. A list of my publications and curriculum vitae are attached as Exhibit A.

7. I was retained in this case to serve as an expert for Burst.com, Inc. I have reviewed Burst's four patents and the various documents submitted by Apple to the Court. I have also read and studied the Court's claim construction ruling in this case. I have also read and studied Apple's summary judgment briefs on invalidity, including Apple's first motion, its reply brief and its second motion.

8. 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.

BUBBLE MEMORY VS RANDOM ACCESS AND IMMEDIATE RECALL

9. The phrase “**random access storage**” refers to the ability to store a chunk of data (e.g. a word of data) in a particular location in memory and retrieve that chunk from memory, where such a location is normally identified by its address. Such a chunk could consist of a small amount of data, e.g., 2 or 3 bytes, depending on the processor and the memory design but it is always a tiny fraction of the total number of bits that the memory can hold. Consider the analogy of a random access storage to a large cabinet of bins containing parts, such as screws and nuts, where there is a large number of small storage bins, each labeled with a unique address. Random Access Storage means you can place an item in any particular selected storage bin having a specified address. (In other

words, you are able to choose any arbitrary location.) Retrieval of data from a random access storage is done by specifying the address of a particular location.

10. The magnetic bubble memory (MBM), invented in 1967 at AT&T Bell Laboratories, in its basic form does not have the feature of random access storage as described above. In later years, modified versions of bubble memory were developed that offered a form of random access storage with the use of a major-minor loop architecture. In general for an MBM, the data is inserted serially one bit after another until all the desired data to be stored has been entered. Furthermore, the memory has the character of a circular shift register. Consider the analogy between the way a dry cleaner shop stores and retrieves a suit from a circular rack containing hundreds of cleaned clothing items. The clerk pushes a button and the clothing items move around a circular track until the desired item reaches the clerk's location. Then the button is released, the rack stops moving, and the clothing item is retrieved. An alternative analogy is a ferris wheel at an amusement park. To load or unload a particular passenger, the operator must rotate the wheel until the desired car is at the load/unload location. Of course these mechanical analogies are neither instantaneous nor fast, but in the case of a bubble memory, the time it takes for the desired data to reach the load/unload location could be a matter of milliseconds and, depending on one's subjective criterion for speed, it may be thought of as offering "immediate recall".

11. I believe that Apple (Second Motion 11:28-12:5) has misinterpreted Kramer's words "immediate recall of the data in any portion of the memory". First of all, Kramer mentions this only in connection with prior art and not in connection with his invention, so he is not claiming that the memory in his invention offers immediate recall. Furthermore, in describing his invention, Kramer clearly refers to the fact that the data has the character of a circular shift register

> The memory 22 is preferably organised so as to appear to be a circular shift register of the required size and is clocked at the same speed, controlled by the memory control clock 36, during recording and replay. One "bit" is presented to the memory at a time. (Kramer 4:1-4)

which makes it abundantly clear that he is not using one of the more sophisticated MBM architectures with random access but rather he is indeed proposing the use of a conventional MBM without random access.

12. In fact, the single loop MBM has the simplest design and easiest interfacing compared to other MBM architectures. The data moves around the circle rapidly so that any piece of data can be recalled when it comes around the circle to the access point. Since Kramer explicitly proposed bubble memories, his mention of "immediate recall" does not suggest or imply that a random access memory could or should be used. (Kramer does not use the words "random" or "random access" anywhere in the patent.) In fact, no memories have literally "immediate recall". There is always some time interval (e.g., measured in milliseconds or microseconds) needed to store or retrieve a unit of data with any kind of digital memory, whether it be an MBM, a semiconductor RAM or any other kind of memory.

13. Kramer specifically teaches the advantage of the <u>serial</u> character of the MBM, by noting that

> All the illustrated links (connectors) are serial, to minimize the number of links needed. (Kramer 3:3-4)

One with ordinary skill in the art would recognize that the alternative to serial links are parallel links (e.g., a bus of 16 parallel lines to simultaneously transfer 16 bits from one place to another) and that the complexity of a circuit with parallel links would be much greater than that of Kramer's invention. This is particularly important to help achieve a small portable card. Since an MBM is by nature a serial memory, it is a natural choice for use with serial links in the circuitry, simplifying the system design, implementation, and operation. Kramer further confirms this by noting that

> One "bit" is presented to the memory at a time. (Kramer 4:5).

Kramer is teaching the advantage of using an MBM with a serial circular shift register structure rather than random access storage since the latter requires parallel data transfer. It also has a lower complexity and simpler interfacing than random access MBMs. This choice is further motivated by the fact that the invention is to be used for the sole purpose of playback of an audio work, which by its nature is produced from a serial bit stream. Kramer never considers editing of audio as a part of his invention (and Apple has not cited any logical reason for editing a song by a user of Kramer's portable card meant for listening to the song).

14. Note also that MBMs had the advantage of nonvolatility unlike conventional semiconductor random access memories (RAMs). (Nonvolatility means that the stored work will be preserved even when no power is applied to the memory.)

15. Based on the above, one of ordinary skill would not have an apparent reason to use random access storage in Kramer and it would not be obvious that any benefit could be gained by doing so. In fact, it would increase cost and complexity.

KRAMER DOES NOT TEACH FASTER THAN REAL TIME

16. The Introduction of Apple's reply brief (dated June 21 , 2007) states:

> The Kramer patent discloses sending compressed audio data between memory cards at "a speed much faster … than that required for actual sound reproduction," so that those cards can be recorded in a "very short time" (reply brief 1:3-5)

Apple is incorrect in several ways:

(a) Kramer does not teach the use of compressed data.

(b) Nowhere does Kramer disclose sending data faster than real time between memory cards.

(c) Kramer's "speed much faster…" statement does not imply faster than real time.

(d) Recording in a "very short time" does not imply faster than real time.

17. The crux of Apple's argument is their claim that the data output from memory card to demultiplexer for producing analog audio output is faster than real time and consequently they infer that the same alleged faster than real time rate is used for sending the data in digital form to a second (replay) card. I will provide sufficient detail below to make it absolutely clear that Apple is wrong.

Memory Output Is Not Faster Than Real Time

18. Apple (reply brief 6:27 to 7:7) is trying again to say that the speed of the data from memory to the decoders is at least 100 times faster than real time, which if it were true would imply (using my own statement) that the direct data transfer from memory card to another card is necessarily also at this speed. I believe Apple has incorrectly understood Kramer and that one with

ordinary skill in the art would realize that Apple's argument is incorrect. In particular, an understanding of the operation of sub-band coding is needed here.

19. Let me restate the black box argument (given in Paragraph 13 of my first declaration) in a different way so it is crystal clear. Suppose that the data from the memory were fed one time into the black box at 100 times faster than real time and this data represents a piece of music of duration 200 seconds. Once the data has been sent to the box there is no additional input to the box. Then it would take 2 seconds for all of the data in memory to pass into the black box. But the box produces real time music output that lasts for 200 seconds rather than a speeded up chipmunk version of the entire song. So after the first 2 seconds, there will remain another 198 seconds of music to come out of the box during which time no data is coming into that box. Where would this music come from if there is no memory in the box and no input to the box? Only Harry Potter could create such a box. Hence, as long as the data is transmitted only once from memory to the box, it is necessary for the data to be sent in real-time, taking 200 seconds to be delivered to the box, in order to produce the real-time audio output. Note that this argument is entirely independent of the specific encoding method used for the digital representation of the audio.

20. Apple, in trying to refute my black box statement, says:

> The fact that "the multiplexer, decoders and analog signal mixer" have no memory does *not* imply that the decoders must receive the sound data at a real time rate. [Apple reply brief 7:18-20]

But I did not make this implication. You could put data into the box at any rate you want, however, if the output of the black box is real time audio (which Kramer states), then and only then can one conclude that the absence of memory in the box implies the input data must be at a real time rate.

Apple's Theory of 100x Repetition of Memory Output is Nonsensical

21. Apple has also attempted to circumvent my black box argument by saying that the entire data from the memory is repeatedly sent to the box 100 times during a play back period. This is a very bizarre proposal! In my entire career, I have never encountered any case where the memory is repeatedly read out multiple times in order to achieve a one-time playback of a single audio work. It

is bizarre because anyone with ordinary skill in the art knows that the cost and complexity of a circuit becomes larger as the clock speed increases, so one would always want to use the lowest clock rate necessary. To increase the clock rate 100 fold is a drastic change and would almost certainly exceed the technological limitations of bubble memories. It is clearly easier to send the data only once at the lower clock rate. All the information needed to reproduce the audio for a human listener is available from a single read out of the memory. In attempting to justify their erroneous interpretation, Apple points to Kramer's words

> The decoder can read the data at the required slower reproduction rate by taking, *e.g.*, only one out of every 100 bits of information presented to it at a time; the intervening 99 bits will be read on subsequent cycles of memory (Kramer: 4: 47-51)

and claims that this means that the entire contents of the memory is read out 100 times and that each "cycle of memory" means one complete reading out of the entire memory contents. So, in the context of my example above of a 200 second audio work, they are saying that the data from the memory comes out in 2 seconds for each "cycle" and the same data is being fed out repeatedly 100 times during a play back period. Hence, Apple is arguing that each "cycle" of data transmission would be 100 times faster than real time, but by feeding it out repeatedly for 100 times, they are able to produce real-time output.

22. I will explain below why the 100x repetition theory is incorrect, but for the moment suppose that it were correct. Under this hypothesis, it would take 200 seconds to repeatedly feed out the data 100 times from the memory for a 200 second song. Hence, this transfer from memory would not really be faster than real-time and if the same 200 seconds were used to send this repeated data directly to a replay card, the card to card transmission would require real time to send the 200 seconds of music. Kramer does not state anywhere that the memory is instructed differently by the controller for the two options of (i) analog output from the memory card or (ii) data transfer to another card. I realize Kramer discloses a 3.5 minute or 210 second song, but for purposes of simplicity in this declaration, I have used 200 seconds throughout because it is a simpler number.

23. It is highly unlikely that Kramer teaches sending the entire memory data repeatedly 100 times in order to play one song for the following reasons:

(a) For decoding audio on the memory card, it does not make sense to transfer data faster than needed by the demultiplexers and decoders. In fact, it would require special synchronization for the demultiplexer to be able to select the correct one bit out of 100 arriving bits and also the card would have to generate a secondary clock signal that is 1/100 the speed of the main clock signal. It would be far simpler to simply read out the data at the much lower clock signal as actually needed.

(b) The term "cycle of memory" used by Kramer in the above cited quote does not refer to the time interval for reading the entire data from memory but rather to the time between reading successive units of memory (e.g., a single bit from storage). In the context of MBMs, a memory cycle means the time for a single shift of the data one position forward on the circular shift register.

(c) Kramer has implicitly suggested a data rate of 305 kb/s to represent the audio (based on 8 Mbytes needed to store 3.5 minutes of audio). To extract the data for real time reproduction a clock rate of 305 kHz would be suitable. To extract the data at 100 times real time, a clock rate of 30.5 MHz would then be needed. However, in extensively reviewing the literature on MBMs, the highest operating rate for a MBM that I could find is only 2 MHz. Thus, the needed rate greatly exceeds the known limits of MBM operation.

24. There is indeed a logical explanation for Kramer's phrase "one out of 100 bits" that does not support the theory of 100x repetition of the memory output. This requires an understanding of sub-band coding. Let us examine the context in which the phrase "one out of 100 bits" appears:

> The demultiplexer 24 directs digital data to the appropriate decoder 26 in accordance with the sub-band of the data. The decoder can read the data at the required slower reproduction rate by taking, e.g. only one out of every 100 bits of information presented to it at a time; the intervening 99 bits will be read on subsequent cycles of memory, so the data in the memory is held in an interleaved fashion. (Kramer 4:45-52)

25. The demultiplexer operation and the word "interleaving" are the keys to understanding what is being done here. A sub-band coder encodes an analog audio source with multiple encoders, each producing a separate bit stream and the bits in these streams are then interleaved to produce a single composite bit stream representing the audio signal. In order to decode the audio, the successive bits in the bit stream coming from memory are deinterleaved by the demultiplexer, *i.e.*, they are directed to successive sub-band decoders. Demultiplexing this composite bit stream is analogous to a single waiting line at a large bank where a clerk directs successive customers to different tellers. (Imagine that there are 100 tellers.) The clerk serves the role of demultiplexer. In each "cycle of memory" one bit is read out by the memory and directed by the demultiplexer to one decoder. In subsequent cycles, the demultiplexer will receive successive 99 bits and direct these bits to 99 other decoders. Then the process repeats with the 101st bit going to the first decoder, and so on. Thus, Kramer's words above make sense for the case of 100 sub-bands with 100 decoders. In other words, the data in memory is held in an interleaved fashion. And clearly the interleaving is so that the demultiplexer can route bit 1 to decoder 1, bit 2 to decoder 2, etc. up to bit 100 to decoder 100. Each of 100 decoders is fed 1 bit that belongs to it and then the process repeats with the next 100 bits from memory. This is consistent with multiplexing practice and with the usual meaning of a memory cycle. The only logical reason for holding the data in memory in an interleaved fashion is that the data created by the encoding system for the various sub-bands must be interleaved, e.g. multiplexed, into one serial data stream as needed for transmission or reception of sub-band coded audio. Kramer does not give any indication of any other reason or function for the example of reading one out of every 100 bits.

26. The plain language of Kramer supports the conclusion that there are 100 sub-bands. Kramer says:

> The number of frequency bands may correspond to the number of instruments/voices in an ensemble, and should correspond to the number of decoders 26. (Kramer 3:27-30)

If we take this statement at face value (although Kramer is wrong here; he evidently does not fully understand sub-band coding), the number of sub-bands should be large enough to handle the most demanding of musical works. It is well-known that there are around 80 to 100 instruments in a symphony orchestra, so the choice of 100 sub-bands is consistent with Kramer's reasoning. In fact, the number 32 of sub-bands is quite common and occasionally 64 might be used, but 100 would be unusually large. Nevertheless, it is a reasonable consequence of Kramer's understanding of the number of bands needed and of his apparent intent to allow for all possible kinds of music. This in combination with Kramer's interleaving statement leads to the most logical conclusion that there are 100 sub-bands.

27. Apple in its reply brief (8:20-24) argues:

> Kramer expressly says that the memory output is "at a speed much faster (at least 100 times) than that required for actual sound reproduction." "Actual sound reproduction" requires reproduction of *all* of the sub-bands in a signal, not just one of them, so this statement shows that the data for *all* of the sub-bands is output "at a speed much faster" than real time.

In fact, *actual sound reproduction* does not require reproduction of all of the sub-bands. A sound can be produced from a single sub-band signal even though it contains only a limited range of acoustic frequencies and is not aesthetically pleasing or useful by itself when compared to the full musical work. One may still hear the beat or rhythm of the music when listening to the audio reproduced from a single sub-band. A few sub-bands can together produce sound reproduction that will sound very similar to the original, although not with the full quality obtained when all of the sub-bands are combined. Apple's last phrase above "so this statement shows …" is their interpretation of Kramer and not based on fact. The deinterleaved data going to each sub-band decoder represents by itself a real-time audio signal that is one component of the entire audio signal. There are 100 such sub-band audio signals, each of duration 200 seconds and each of which is in effect a separate real time audio signal so that jointly the data describes 100 audio signals each of 200 seconds duration. Consequently, this data would take 20,000 seconds (200 x 100) if they were sent sequentially rather than concurrently. But since the interleaved data representing all of them has a 200 second duration, it is natural for Kramer to describe the memory output as being at a speed much faster than that required for actual sound reproduction. This is simply part of the character of sub-band coding: each audio signal is reconstructed by combining (in the analog signal mixer 28 in Fig. 1 of Kramer) a multiplicity of sub-band signals to produce a single final audio reproduction. Furthermore, Kramer's statement here is consistent with other statements in Kramer that support the 100 sub-band embodiment.

28. Regarding Apple's footnote 13 (reply brief pp. 7, 8):

> As Kramer goes on to explain, in order to accommodate the fact that each decoder ignores 99 out of every 100 bits it receives, the memory stores the audio data in "an interleaved fashion." Kramer at 4:45-52. This interleaving is not related to the supposed 100-fold sub-band encoding Dr. Gersho describes, it is related to the fact that Kramer's decoders [use] only 1 out of every 100 bits it receives in order to achieve the "required slower reproduction rate."

Apple's interpretation of the purpose of interleaving does not make sense to me. First of all, it is not clear how using one out of 100 bits is related to interleaving. Also, if they are suggesting that the decoders *collectively* use only 1 out of the 100 bits that they *collectively* receive, and they rely on repeated transmission of the entire memory contents to ultimately receive all needed data, this would be extremely weird, with a grossly excessive and unneeded complexity and would require at least a very tricky clocking and control mechanism to get the right bits to the right decoder at the right time and even then it is unclear how they can successfully decode the audio. The statement is also ambiguous: To what does "it" refer in the phrase "it receives"? (Possibly "it" refers to "interleaving" or the "decoders" or perhaps "it" refers to one of the decoders.) If the intervening 99 bits are all to be applied to the same decoder at successive memory cycles, then what is being interleaved with what? The intervening 99 bits are not being discarded but are read at subsequent cycles of memory. Does this mean that they are read by the same decoder at later times? I am unable to see any way to make sense out of Apple's interpretation.

Apple's Four Sub-Band Conjecture

29. Apple argues that there are only four sub-band decoders based on the illustrative drawing in Fig. 1 of Kramer. Clearly Kramer showed four only for purposes of illustration. It would not be practical to draw 100 decoders. Nowhere in Kramer's text is there any indication that four bands are used. I have attempted to make sense of this hypothesis without success and can offer the following *reductio ad absurdum* argument to show that Apple is wrong.

30. Suppose there are 4 sub-bands and 4 decoders as Apple contends. Then let's label the bands as R, B, G and P and label the ordered sequence of bits needed by Decoder R as

r1, r2, r3, r4, r5, r6, ….

and the ordered sequence for Decoder B as

b1, b2, b3, b4, b5, b6, …

and similarly for Decoder G

g1, g2, g3, g4, g5, g6, …

and for Decoder P

p1, p2, p3, p4, p5, p6, ….

31. Each of these bit streams is the digital representation of a real time audio sub-band signal. I now examine the implications of Kramer's statement:

> The demultiplexer 24 directs digital data to the appropriate decoder 26 in accordance with the sub-band of the data. The decoder can read the data at the required slower reproduction rate by taking, e.g. only one out of every 100 bits of information presented to it at a time; the

> intervening 99 bits will be read on subsequent cycles of memory, so the data in the memory is held in an interleaved fashion. (Kramer 4:45-52)

As Kramer notes, "the data in the memory is held in an interleaved fashion". In fact it is correct that in sub-band coding, the encoder multiplexes (e.g., interleaves) the data coming from each sub-band encoder to form a serial bit stream[1]. Hence the serial bit that would be produced by the 4-band sub-band encoder, recorded in memory and finally retrieved from memory and applied to the demultiplexer is the sequence of bits:

r1, b1, g1, p1, r2, b2, g2, p2, r3, b3, g3, p3, …. (*)

which continues until the end of the audio work. Hence, according to Kramer's statement above, the appropriate decoder (one of the alleged four decoders, R, B , G, and P so let's say it is decoder R) reads the data "by taking, e.g. only one out of every 100 bits of information presented to it at a time;" so that the sequence of bits it accepts is:

r1, r26, r51, r76, r101, ...

where it should be noted that r26 is the 101th bit in the sequence (*) sent to the demultiplexer, r51 is the 201th bit, etc. But this sequence is useless for reconstructing the first 2 seconds of sub-band R of the audio work (of total duration 200 seconds) since it is missing 24 out of 25 of the needed bits and is grossly inconsistent with the correct sequence that is needed, namely, r1, r2, r3, r4, r5, r6, …. .

32. But under Apple's contention that one cycle is the entire memory readout, Kramer's phrase: "the intervening 99 bits will be read on subsequent cycles of memory", implies that on the second cycle, another one out of 100 bits (one of the 99 intervening bits) will be read by each decoder. In this case, the sequence of bits accepted by decoder R during the 2nd cycle would be

r2, r27, r52, r77, r102, …

[1] The interleaving can be bit interleaving or word interleaving, where a word refers to a group of bits, e.g. 16 bits, which represent one sample of the audio sub-band signal. Kramer refers to a decoder taking one bit at a time, so he is implying bit interleaving.

which is useless for reconstructing the next 2 seconds of sub-band R of the audio work since it is missing 24 out of 25 of the needed bits. Similarly, in the 3rd cycle the sequence would be

r3, r28, r53, r78, r103, …

which is again useless for reconstructing the third 2 seconds of sub-band R of the audio work. The same problem arises with each of the other three decoders. Finally, concatenating the sequences that are produced in the successive cycles in this manner, we obtain:

r1, r26, r51, r76, r101, ... r2, r27, r52, r77, r102, … r3, r28, r53, r78, r103, …

which is grossly inconsistent with the correct sequence that is needed, namely,

r1, r2, r3, r4, r5, r6, r7, ...

So under Apple's contention, this scheme would not produce an intelligible audio signal output from the memory card (or from the replay card if the same data stream were transmitted to a replay card).

RECORDING NOT FASTER THAN REAL TIME

33. Apple uses Kramer's statement

> The memory 22 … is clocked at the same speed, controlled by the memory control clock 36, during recording and replay. (Kramer 4:1-4)

to infer that recording of data is faster than real time since it has the same speed as the data output from memory during replay. But I have shown that the latter is not faster than real time and therefore the recording is also not faster than real time.

CARD TO CARD DATA TRANSFER IS NOT FASTER THAN REAL TIME

34. Apple argues that direct digital output from the memory card to another memory card is at the same speed as the output from the memory card to the demultiplexer and hence the card to card transfer rate is faster than real time. But in fact the output to the demultiplexer is not faster than real time, as I have shown, so the card to card transfer is not faster than real time.

35. Apple cites Kramer's statement that recording can take a very short time:

> When recording is completed, which can take a very short time, the card is removed from the input recorder and can be stored or transported as required. [Kramer 4:6-8.]

and implies that this supports their contention that faster than real time card to card transfer takes place. But the word "short" is rather intangible and begs the question "relative to what?" The best answer to this is that it is short relative to what one might expect. Recording audio data to the memory was not a consumer operation but one presumably to be performed by a music store or other vendor of portable audio memory cards each loaded with a song. One could think of the recording service as in a certain sense similar to a factory's production process or as a service provider (such as a film development laboratory or a laundry service). The delivery of such products or services could require a matter of days rather than minutes and it is quite possible that a consumer might expect a similarly long time for providing a desired audio work on a customer's memory card. So at the time of Kramer's patent, it would have been reasonable to say a "short time" if the recording process would take not days, but only hours or many minutes or even be achieved in real-time. Thus "short" does not imply that the transfer rate of data from the input encoder is faster than real-time.

FASTER THAN REAL TIME NOT OBVIOUS FROM KRAMER

36. Apple (2nd motion 11:25-27) argues that it is obvious to use faster than real time transmission

because of Kramer's statement:

> This output will be at a speed much faster (at least 100 times) than that required for actual sound reproduction. (Kramer 4:24-26)

I disagree. First of all, Kramer's statement does not describe faster than real time transmission. One with ordinary skill in art would understand that Kramer is describing only how data is sent from the memory to the demultiplexer and furthermore that the real time transfer of interleaved audio data for each of a large number of sub-bands each of which is in real time, requires that the data rate be faster than the normal rate needed to reproduce a single audio signal in real time. The higher rate is needed to simultaneously carry the data for multiple audio signals in real time. Therefore there is no burst mode of operation since the data requires 200 seconds to be transferred from memory to the decoders in order to generate real time reproduction for a 200 second song.

37. One with ordinary skill in the art in the late 1980s (or later) would not want to use a higher clock rate than necessary. Furthermore, increasing the clock rate substantially to achieve 100x real time (as Apple suggests), would require going (for example) from a 250 KHz clock rate to a 25 MHz clock rate. While 250 KHz was a reasonable clock rate for MBMs, I have never encountered reports of MBMs operating at clock rates higher than 2 MHz and believe that 25 MHz is not possible. Even if MBMs were replaced with other memory technologies, the issue of clock rates and circuit complexity still makes it unreasonable to use a burst mode of operation that is significantly faster than real time.

38. Not only does Kramer not suggest faster than real time transmission, it would have been entirely surprising and unexpected for anyone to conceive of or propose the idea of transmitting audio in a burst mode given the state of technology and the consumer product market in the late 1980s and early 1990s. Digital audio was used largely for storage although some digital transmission of medium band audio (*e.g.*, FM radio quality) was of interest in the 1980s. In 1992 the famous MP3 audio compression algorithm was finalized as a standard and offered efficient compressed representation of audio. Even then, memory capacities were very small. I believe that the mindset of those working in the field in the 1980s was that digital audio was primarily for storage with factory generated CDs offering a better way to store music than audio cassette tapes or vinyl records. It was also recognized that it could be also used for real time transmission. Flash memories, introduced in the 1980s had very small capacities but grew with Moore's law to become useful in the 1990s for storing multiple songs. I believe the technology and the market were not ripe for burst mode transmission of audio and considerable imagination would have been needed to think that in the future small digital memories would hold hundreds or thousands of songs in which case a rapid way to transfer these songs would one day be desired. (Note that Kramer's proposed card would hold only one or two songs.) Thus I believe there was no motivation and no apparent reason in the late 80s or early 90s to take Kramer's work and improve it by adding faster than real time transmission together with the other elements in Lang.

NO COMPRESSION

39. Kramer is clear and unambiguous in teaching the use of DPCM as a method of converting analog audio to digital form. Nowhere in Kramer is there any description of compression or compression means. Dr. Wicker, in his declaration of July 13, 2007 repeatedly and incorrectly states otherwise when he says:

> "Kramer describes and illustrates in its figures a portable card used to store compressed music." (Wicker ¶13),

> "Kramer also describes compression of the audio information … " (Wicker ¶13),

> and

> "… to use the compression, storage, and outputs described in Kramer …" (Wicker ¶13).

Kramer never describes compression, never uses the words "compressed" or "compression" and never mentions bit rate reduction. On the other hand, Kramer says:

> "The music signal is encoded (outside the illustrated system), into digital form, by any suitable technique; that known as differential pulse code modulation (DPCM) is suitable." (Kramer 3:9-12).

Here Kramer is clearly saying that the conversion from analog into digital form is suitably performed by DPCM. Kramer goes on to say:

> A pulse code modulation coder quantises sampled sound amplitudes; the differential technique is more efficient and utilises the redundancies present in the sound, the change in analogue signal is recorded digitally … (Kramer 3:14-18)

This leaves no doubt that he is describing the differential encoding technique, DPCM, as a more efficient way to go from an analog signal to a digital one than the pulse code modulation (PCM) technique. Kramer also uses the phrase "digitally encoded" (1:59), "encoded analogue signal" (2:21), and "an analogue signal which has been encoded into digital form" (2:11-12) all of which show that Kramer was focusing on efficiently converting a signal from analog to digital rather than reducing the bit rate of a digital signal (compression).

40. Apple incorrectly refers to "Kramer's silence on whether to use analog or digital differencing…" (reply brief 5:5-8), where "analog differencing" is used in direct encoding of analog to DPCM digital form and "digital differencing" is used in bit rate reduction with a digital implementation of DPCM. But Kramer specifically says:

> The music signal is encoded (outside the illustrated system), into digital form, by any suitable technique; that known as differential pulse code modulation (DPCM) is suitable.

The use of the phrase "encoded .. into digital form" with DPCM being a suitable technique clearly shows he is talking about techniques for digitizing the signal, i.e., getting the signal into digital form and he cites DPCM as a suitable technique for this purpose. This clearly shows that he is not silent on this matter.

41. Apple's use of the phrases "analog differencing" and "digital differencing" is clearly intended to give the impression that there is minimal difference between the use of DPCM for encoding an analog signal and the use of DPCM for reducing the bit rate of a digital signal. But there is indeed a fundamental difference between these two techniques, not only in how the differences are formed but also in the feedback loop and the very different technology used in implementing the two distinctly different techniques. DPCM for analog to digital conversion is implemented with analog circuitry and bit rate reduction via DPCM is implemented entirely with digital circuitry. The former uses analog electronics with such components as diodes, capacitors and resistors and voltages and currents have a continuous range of amplitude values while the latter uses logic circuitry where binary numbers are manipulated.

42. Apple's argument that Kramer discloses compression says

> "Finally, the fact that Kramer discloses compression is evident from the language of the reference itself, and not just from its reference to DPCM. Kramer states that "an 8 megabyte memory 22 should allow recording of at least 3 1/2 minutes of music." 3:35-37. That statement alone shows that the audio is compressed" (Apple reply brief 5:9-12)

This conclusion is false. By noting that DPCM can allow music to be more efficiently encoded at this relatively low bit rate does not imply that bit rate reduction was used, but only that an efficient way of getting audio into the digital world can be beneficially used in comparison with simple quantization or pulse code modulation (PCM). Kramer himself states this:

> (A pulse code modulation coder quantises sampled sound amplitudes; the differential technique is more efficient …)"

Quantization is purely analog to digital conversion and comparing its efficiency with the differential technique is further confirming that Kramer is viewing the differential technique as a preferred analog to digital conversion method.

43. Apple's quote of Dr. Hemami's statement is used to justify the false conclusion. She said:

> "a digitized uncompressed wideband audio signal" uses approximately 18.5 megabytes to store 3 1/2 minutes of music.

But her statement is simply saying that audio encoded into digital form (*e.g.*, with PCM) without subsequent bit rate reduction has a very high bit rate. Apple goes on trying to confuse the issue by then quoting Kramer and suggesting but not justifying that it supports their argument. Of course DPCM is more efficient than PCM, it goes from analog to digital in a more efficient way because it produces a lower bit rate representation than does PCM. Comparing the two techniques shows that the DPCM has a substantial saving in information.

44. Apple argues that it would be obvious from Kramer to use a digital version of DPCM wherein compression takes place instead of the ordinary use of DPCM as an analog to digital encoding technique. For this argument, Apple (2nd motion brief 11:13-15) cites from the Bellamy book that is focused on telephony that digital signal processing "is generally the most effective means of implementing a DPCM algorithm" to perform DPCM as a way to encode analog signals. But Apple either is unaware or overlooks the technical and historical context, in which this remark was made, leading to a deceptive and false impression. As I will explain below, use of an analog to digital conversion technique (other than DPCM) followed by a digital version of DPCM for bit rate reduction would replace the simple technique of DPCM for digital encoding as used in Kramer with one that is more costly and complex to implement and would not yield any performance advantage in the context of Kramer.

45. DPCM, invented by C.C. Cutler (U.S. Patent 2605361), was based on analog circuitry and offered an alternative to standard analog to digital converters for digitizing an analog signal. When digital telephony was first introduced in the 1960s, PCM (rather than DPCM) was selected as a simple and effective way to represent voice in digital form and this representation remains the dominant standard in the telephone industry today. In the 1970s, motivated by the need to squeeze additional voice circuits over limited capacity satellite and submarine cable links, engineers needed a way to convert digital voice signals from the standard PCM representation to a more efficient lower bit rate digital representation of voice when terrestrial PCM trunk circuits arrive at a satellite or submarine cable gateway station. These terrestrial circuits were already carrying speech that had already been converted into PCM digital format at a different geographical location so analog to digital conversion was not needed at the gateways. Consequently, transcoders were developed to convert the PCM format to DPCM and reduce the bit rate from 64 kb/s to 32 kb/s. Digital circuitry was used to convert from one digital representation of voice to the other. Hence much effort went into the design of digital transcoder chips that were needed in large quantities at these gateways. In this context, conventional analog to digital encoding with DPCM was not applicable since the arriving signals were already in digital form. The alternative, to convert the standard digital PCM format back to analog voice and then re-encode the analog with DPCM was not an acceptable solution since it would degrade the voice quality and add complexity.

46. Apple’s citation (from a telephony book) that digital signal processing is more effective for implementing a DPCM algorithm is inappropriate and inapplicable. Unlike telephony, where the need for transcoders cannot be avoided, Kramer’s invention has no need of transcoders since the analog source is directly available for encoding into the digital form to be used for storage in memory. Furthermore, the cost and complexity of digital signal processing is far greater for processing wideband audio than it is for telephone speech due to (a) the higher clock speeds required for audio compared to speeds needed for telephone speech and (b) the number of bits of resolution needed to quantize audio compared to speech. Speech signals are sampled at 8,000 times per second and each

sample is quantized with 8 bits whereas wideband audio signals (as in a CD) are sampled at 44,100 times per second and quantized with 16 bits. This is an enormous difference and mid 1980s technology for digital signal processing was very limited in the number of elementary operations per second that it could perform. To replace the DPCM encoding of wideband audio with a high resolution analog to digital converter and followed by a digital implementation of DPCM operating at the needed clock rate would not make sense. The cost and complexity of the converter alone would much exceed the cost of the DPCM encoder used by Kramer. Hence it is not obvious to replace the DPCM encoder with analog to digital conversion and digital DPCM.

47. Apple (reply brief 4-18-19) tries to argue that DPCM is compression by citing this quote from Bellamy:

> Since the range of sample differences is less than the range of individual samples, fewer bits are needed to encode difference samples.

This does not support the conclusion that DPCM implies compression and neither does the quote cited in Apple's footnote 9 of the same page. In these quotes, Bellamy is simply comparing the effectiveness of DPCM over PCM as alternative ways to convert a source from analog to digital. These quotes do not in any way support the argument that DPCM entails bit rate reduction. Similarly, Apple's quote (reply brief 4:19-22) from Jayant's introduction to his Chapter 6 refers to reductions in bit rate compared with encoding without the differential feature. In fact Jayant's introduction is followed immediately by a section entitled "DPCM versus PCM: The Prediction Gain" which discusses the benefits of encoding with the differencing feature compared to encoding with PCM. Again, this quote does not support the argument that DPCM implies bit rate reduction.

NO SINGLE HOUSING

48. The apparatus performing the digitization described in Kramer was not described as nor expected to be portable. Hence, in order to avoid the much higher costs of miniaturization it would be physically much larger in size with relatively large circuit boards compared to the size of the portable cards. The portable cards would not be embedded inside a metal housing containing the

digitizing equipment since they are meant to be portable. The likely scenario is that the storage card would be connected to the digitization equipment by a cable or possibly be plugged into the metal housing, similar to the way a household appliance is plugged into an electrical outlet rather than be embedded in the same housing.

49. Kramer does not describe any transceiver apparatus. His memory card does not have encoding capability, nor is there any apparent reason to give it such a capability. His encoding system does not specify an input, random access storage, or FTRT output. The missing elements of a transceiver apparatus are not inherent in Kramer. Also given the state of technology and the consumer market in the late 1980s or early 1990s, I see no motivation or reason for one to modify Kramer's memory card or his vaguely specified encoding system to become a transceiver apparatus as claimed in Lang. Therefore it is not obvious to add the missing elements in Kramer to create a transceiver apparatus.

COMPUTER GENERATED AUDIO

50. Apple incorrectly states that CompuSonics expressly describes computer generated information Apple (24:15-21). Apple cites the paragraph on stretching a sound in the CompuSonics brochure. Dr. Wicker states that the 'stretching' operation creates audio/video information generated by a computer (Wicker ¶ 26). First of all, the paragraph does not mention 'computer generated'. Secondly, CompuSonics does not describe the generation of a sound by computer nor does it mention the term 'computer generated' but rather it describes taking some original sound bite (previously recorded and stored) and manipulating it by extending it in a tape loop almost indefinitely. One with ordinary skill in the art knows that computer generated sounds refer to synthetic sounds that do not originate from a recorded acoustical signal. In the 1980s and earlier, computer generated sounds were well known (notably computer music that did not originate from any man-made instrument). Apple is untenably stretching the definition of computer generated sound.

51. I see no apparent reason why computer generated audio/video information would be combined with the elements disclosed in Kepley, Walter, Gremillet, and Tescher, or Kramer. The example cited by Dr. Wicker in paragraph 26 is not adequately explained, but perhaps he is referring to use of a digital time stamp added to a voice mail message that can be displayed on the listener's device. In this case, it would be computer generated data and not computer generated audio/video information. Hence it is not obvious to combine computer generated audio/video with the elements of the cited references.

SILENCE COMPRESSION

52. Apple makes the bizarre statement:

> Since silence is the absence of sound, silence compression necessarily compares consecutive samples and codes the differences as the number of samples that do not contain changing signals. (Apple's reply brief 15: 1-3)

This is totally wrong. I have never heard of anyone in the field doing a comparison of consecutive samples to decide whether silence or speech is present and if such an attempt were made it would fail. The remaining part of the statement has no clear meaning and does not make sense to me. The "differences" perhaps refers to a collection of difference values, each of which is a difference between two consecutive samples. If so, "codes the differences" would apparently mean 'counts the number of differences that do not contain changing signals'. But the term "changing signals" has no clear meaning. If the intent is to refer to changing values of the successive samples, then virtually all nontrivial signals are forever changing.

53. Furthermore, one with ordinary skill in the art would know that "silence" in the context of silence compression does not mean the absence of sound but rather the absence of speech sounds. During silent intervals the consecutive samples are still changing due to the ever-present background noise and understanding this is essential to the design of silence compression systems. The key part of silence compression is Voice Activity Detection (VAD), based on measuring short term **energy levels** and not on sample differences. In VAD, each interval (e.g., typically 10 to 30 milliseconds) identifies if speech is present or if there is only background noise present.

54. Silence compression is most useful for speech in a two way voice conversation since one speaker is silent about half the time. It is much less of value for voice mail since there are relatively few silent intervals and they are of much shorter duration. It is of no value for music and I have never heard of it being used for any wideband audio systems intended to carry music.

COMPRESSION MEANS - CLAIM CONSTRUCTION ISSUE

55. An issue raised by Apple is the interpretation of the claim construction:

> The language of the Court's construction of "time compressed representation" -- "a version of audio/video source information having a reduced number of bits" -- implies only that the number of bits must be reduced relative to another representation, not that bits must be created and then discarded. (reply brief 4:11-14).

In my view, the phrase "having a reduced number of bits" means "having a number of bits that has been reduced from the number of bits of a prior form of the representation of this source." The word "reduced" is an action verb and suggests that an action was taken to lower the number of bits needed to represent the audio/video source information. Apple's interpretation is using the word "reduced" in the intangible way that commercials say a detergent makes clothes "cleaner" without specifying relative to what.

I declare under penalty of perjury that the foregoing is true and correct.

Executed August 28, 2007

/s/ Allen Gersho
Allen Gersho