



Neutral Citation Number: [2021] EWHC 1377 (Pat)

Case No: HP-2020-000020

IN THE HIGH COURT OF JUSTICE
BUSINESS AND PROPERTY COURTS OF ENGLAND AND WALES
PATENTS COURT (ChD)
SHORTER TRIALS SCHEME

The Rolls Building
7 Rolls Buildings
Fetter Lane
London EC4A 1NL

Date: 26th May 2021

Before:

LORD JUSTICE BIRSS
(Remotely via Teams)

Between:

FACEBOOK IRELAND LIMITED (a company incorporated in Ireland)	<u>Claimant</u>
- and -	
VOXER IP LLC (a company incorporated under the laws of the State of Delaware)	<u>Defendant</u>

MARK VANHEGAN QC and JAANI RIORDAN (instructed by **Freshfields Bruckhaus Deringer LLP**) for the **Claimant**

DR. BRIAN NICHOLSON QC and CHRISTOPHER HALL (instructed by **Quinn Emanuel Urquhart & Sullivan UK LLP**) for the **Defendant**

Hearing dates: 12th, 13th 15th April 2021

Covid-19 Protocol: This judgment will be handed down remotely by circulation to the parties' representatives by email, release to BAILII and publication on the Courts and Tribunals Judiciary website. The date and time for hand-down will be deemed to be at 10:30am on 26 May 2021.

Approved Judgment

I direct that pursuant to CPR PD 39A para 6.1 no official shorthand note shall be taken of this Judgment and that copies of this version as handed down may be treated as authentic.

.....

LORD JUSTICE BIRSS

Lord Justice Birss :

1. This case is about European Patent (UK) No. 2 393 259 entitled “Telecommunication and multimedia management method and apparatus”. The defendant Voxer contends that a live broadcast feature offered by the claimant Facebook infringes the patent. Facebook denies infringement and contends the patent is invalid. The proceedings were started by Facebook as an action for revocation. Voxer counterclaimed for infringement. The matter has been tried under the Shorter Trials Scheme.
2. The application for the patent was filed on 29th April 2008 claiming priority from a series of US filings, the earliest of which was made on 28th June 2007. This is the relevant priority date in this case. The patent was granted on 17th August 2016 based on a divisional application.
3. Voxer made an unconditional application to amend the claims and then made a second unconditional application. The latter is the only set of claims now in issue. They are set out in Annex A. The amendments are objected to on added matter and clarity grounds, and on the basis that they fail to cure the invalidity. The Comptroller has made adverse observations about the amendments. Voxer contends that those observations essentially mirror and are based on points taken by Facebook in the litigation. If the amendments are refused entirely then it is common ground the patent must be revoked. It is conceivable that the only amendments which might be refused are to dependent claims, in which case, despite Facebook’s submission to the contrary, I believe the right thing to do in that case would be to invoke the partial validity jurisdiction under the 1977 Act, allow the allowable amendments, refuse those which must be refused and renumber claims accordingly.
4. In terms of independent validity, the focus of the case has been on claim 1. Claim 5 is alleged to be independently valid and infringed. The allowability of the amendments to produce claims 2, 3, 4, and 10 needs to be considered and claim 10 is said to be independently valid but not infringed.
5. Voxer contends that the patent is infringed by the live broadcast feature offered to users via the Facebook website and through the Facebook and Instagram Apps as they operate on iOS devices (i.e. devices sold by Apple). Infringement is advanced on a normal construction of the relevant claims and on the basis of the doctrine of equivalents in two respects. Infringement under s60(2) is also alleged in certain respects. On equivalents, in addition to its denial, Facebook also contends it has a *Formstein* defence.
6. Voxer had claimed there was infringement by the feature offered via the same Apps as they operate on Android devices. However following clarification of how those Android Apps worked shortly before trial, Voxer withdrew the allegation of infringement. I dismissed that infringement claim and granted a declaration of non-infringement.
7. Facebook challenges validity on various grounds. The claims are alleged to lack novelty and/or be obvious over two items of prior art: WO 2006/121550 (Atarius) and US 2006/0003740 A1 (Munje). Both were published before the earlier claimed priority date of the patent. Facebook also alleges the claims are invalid for insufficiency.

8. Facebook called expert evidence from Dr Tim Kindberg. Dr Kindberg is an expert in distributed systems and has 30 years experience as a platform and application developer. His opinions essentially supported Facebook's case. He was a good witness, seeking to help the court with his answers to questions posed in cross-examination. That does not mean I will necessarily accept everything Dr Kindberg said, for example I have not accepted some of his evidence about what was common general knowledge.
9. Voxer called expert evidence from Mr Ashley Unitt. Mr Unitt is a software engineer. In 2000 he co-founded what became a market leading media messaging company and served as its Chief Technology Officer for that company for 16 years. His opinions essentially supported Voxer's case.
10. Facebook submitted that Mr Unitt was argumentative and confused. He was neither of those things. Facebook also submitted that he was at times internally contradictory. The aspect where this point has significance arose from a contrast between his view about what was taught by the patent (which often focussed on conversations) and his broader view about the alleged infringement (which did not). That is a specific issue I have taken into account where it mattered. It is no reason to apply a general discount to his opinions.
11. Contrary to further submissions of Facebook about Mr Unitt:
 - i) The fact he could not recollect where a point of detail (half-duplex) had come from was not sinister. Mr Unitt plainly knew what half-duplex communication was. That was not something suggested to him by anyone else. He did not think his use of it to characterise a particular point had been suggested by the lawyers, but given all the discussions which had gone on, he could not say with 100% certainty.
 - ii) The fact he picked up some patent lawyer speak (such as the phrase "term of art") and may have misunderstood it does not tell the court anything useful. The idea that he can be criticised for expressing the view that something may involve "an inventive step over" a particular item of prior art is absurd. There is a jargon in patent cases which experts inevitably pick up and use.
 - iii) The fact he sometimes said a document had to be read as a whole was unexceptional. There may be a specific point about how he approached Munje prior art but it has no wider significance.
 - iv) His focus on "use cases" was not unfortunate. It was helpful.
 - v) Mr Unitt did not lose sight of his job as an independent witness.
12. In my judgment, contrary to Facebook's submissions, Mr Unitt was in fact a good witness, also seeking to help the court with his answers to questions posed in cross-examination. That does not mean I will necessarily accept everything he said either, but I will address specific issues when they arise in context.

The skilled person

13. The patent is directed to someone (probably a team) concerned with designing and implementing a multimedia communications system. As Dr Kindberg put it, there would be an application programmer with experience in multimedia (voice and video) networking, streaming and messaging. There would also be a back-end developer with knowledge of streaming protocols and responsible for server-side software. I believe in substance these two amount to the same team/person as the one posited by Mr Unitt, who emphasised the need for experience in telecommunications networking, particularly voice over Internet Protocol (VoIP), and the processing and management of media. If there is a difference between them then I prefer Dr Kindberg's formulation because it explicitly highlights the significance of application software (on a mobile device running on a phone or laptop) and of server-side software, and because it explicitly highlights video as well as voice.
14. Dr Kindberg also contended there was a third member of the team, a mobile user interface developer. Mr Unitt did not agree. However, his disagreement came down to a point about the skills of the team. He agreed that the team would build a system with a user interface, but Mr Unitt's conception of the skilled team was one with sufficient skills to build a workable user interface without input from a specialist. I doubt it matters but if it does, then again I prefer Dr Kindberg's formulation. That is because I am quite satisfied that user interface development skills would be required of the skilled team (and did exist).

The common general knowledge

15. The law relating to common general knowledge is well known. There is a specific point about geography arising from an alleged distinction between what might be known in the USA as opposed to the UK. It relates to PTT (below). Separately, Voxer submitted that a fact known only to some skilled people is not common general knowledge and also that one may need to take care not to conflate knowledge of the details of something with knowledge that something existed or was possible. I agree with both of these submissions.
16. There is a great deal of technical background information about telecommunications and standards which is common ground and unnecessary to set out. In this section I address matters of common general knowledge which have a particular bearing on the issues to be decided.
17. One point is the difference in a communications system between applications which employ a peer to peer model and those which apply a client-server model. Consider two mobile phone devices connected to the internet via wireless telecommunications networks. For this purpose one can take all that network infrastructure for granted and ignore it. An application on one device could communicate with a corresponding application on the other device by sending messages to the other device through the network infrastructure without the involvement of any intervening application on the network. This is peer to peer communication. The contrast is with the client-server model in which the applications on the devices each communicate with a special application server which exists somewhere in the network. Each model has advantages and disadvantages. There could be more than one application server on the network. All this is part of the common general knowledge of the skilled team.

18. There are various techniques for communicating between two devices which are part of the common general knowledge. They include telephony, instant messaging, video streaming, and PTT (push to talk).
19. Irrespective of the communications technique used, all systems involve some delay between the speaker speaking into their microphone and that sound being played on the loudspeaker of the receiving device. The skilled team knew that for two people to have what they regard as a natural conversation without appreciable delay, the normal maximum delay or latency must be no more than 300 milliseconds and ideally less than 150 ms. By the priority date the International Telecommunications Union (ITU) had published a graph showing the relationship between delay and user satisfaction in interactive communications (such as a phone call). Below about 200ms very many users are satisfied but by the time the “mouth to ear” delay reaches 500ms nearly all users are dissatisfied.

Telephony and VOIP

20. The only point to mention relating to telephony itself is that by this time VoIP was part of the common general knowledge. The details of the differences between VoIP and the other forms of mobile telephony do not matter.

Instant Messaging

21. The existence of a large number of instant messaging protocols was common general knowledge. The fine detail of individual schemes was not, but the principles of their operation was common general knowledge. Examples are AIM, ICQ and MSN Messenger. The common general knowledge included mobile versions of these systems, included the concept that voice clips could be sent for later review, and that offline messages could be sent which the recipient would download and review the next time they logged in. The relevance of this latter point is that to do this means that the idea of allowing a device to start transmitting a message for another device without establishing that the recipient device is even switched on, or connected to the network, was common general knowledge.
22. There was a dispute about whether the idea of instant messaging client applications supporting voice and video calling, including multi-party group video, was common general knowledge. In my judgment these concepts were common general knowledge. In other words, the skilled team knew that such systems existed and could be implemented. The details of how the particular systems worked would not be part of the common general knowledge, but the team would be able to build a system which provided those functions if they wished to. It would be a lot of work but that is a different point.
23. It was common general knowledge that text-based conversations in IM applications could be threaded into a series of conversations containing individual messages. It was also common general knowledge that messages could be archived and could be searchable and that text, voice and multimedia messages could be stored locally.
24. There was an issue about the recording of voice calls. Mr Unitt’s business experience meant that he had considerable experience of this problem. At one stage his company

(now called Resilient) became the largest business voice mail provider in Europe. In his first report he said the following about recording of voice calls:

“89. [...] the general concept of storing material on a network was well known, but in a context (for example in a call centre) in which only one party was in control of the recording, typically via pre-defined and coarse rules. [...] recording by the user was only really known in the context of locally connecting a recording device to the line.”

[The ellipsis are inserted because this passage of Mr Unitt’s evidence was focussed on the Atarius prior art but the evidence is about common general knowledge.]

25. However, in cross-examination he accepted that the skilled person knew that there were in existence applications for mobiles and laptops for recording voice and video calls. In my judgment the skilled team knew, as a matter of common general knowledge, that such applications existed albeit that the details of how the particular systems worked was not common general knowledge. Again, the team would be able to build such a system which provided those functions if they wished to.
26. Facebook relied on a product called Trillian which had a paid for version called Trillian Pro. The Trillian Pro application (but not Trillian) had a “time travel” function. I do not accept that was common general knowledge. Mr Unitt had never heard of it.

Video streaming

27. Live streaming of video communications was common general knowledge, including in mobile phone networks (3G). Mr Unitt said it had not gained widespread user acceptance by the priority date. I think that is right, but it is not the important issue. I accept Dr Kindberg’s view that video streaming was growing in popularity and was widely known.
28. A feature of live streaming video was the availability of so-called VCR (Video Cassette Recorder) functionality. This is the ability for the user who receives the video stream on their device to pause/play, rewind, fast forward and replay live video streams. This applied to both live and on-demand voice and video streams. The client device could store a local cache of the received content to enable local playback with all this VCR functionality, independently of other viewers. All this was common general knowledge.
29. Mr Unitt, while accepting that techniques implementing VCR functionality were common general knowledge for video and audio streams using a client server model, did not accept that the same techniques were part of the common general knowledge for voice or video calls or video conferencing. In cross-examination Facebook’s counsel put passages from a 2007 textbook entitled “Multimedia Over IP and Wireless Networks” by Mihaela van der Schaar and Philip Chou. These were said to show that this VCR functionality was common general knowledge in the context of voice/video calls and conferencing. Mr Unitt maintained he had not heard of it. In my judgment that aspect was not common general knowledge.

30. Something which was illustrated by the van der Schaar textbook and which was accepted as common general knowledge was the technique of using a CDN (which stands for Content Distribution Network or Content Delivery Network). This is part of the architecture of an application for delivering content to multiple users. Instead of delivering content from a single server to all clients, a CDN amounts to a network of servers in the transport network so that the content is delivered to users over multiple paths through that transport network. It helps overcome loss and delay problems that afflict streaming media. It improves latency, fault tolerance, scalability and load balancing. All this was common general knowledge.

PTT and over cellular (known as 'PoC')

31. PTT stands for Push to Talk. "Walkie talkie" radios work in that way. It is a single channel half duplex system. A user holds down a button on the radio while they are speaking. When they are finished the other person can reply in the same way. The convention of saying the word "over" when the speaker has finished talking comes from this paradigm, to indicate that the single channel is now free.
32. PoC stands for PTT Over Cellular. This is a PTT system built using the mobile phone data network. The idea of using the mobile phone system this way was an old one at the priority date. The term PoC is often being used to refer to a particular protocol for implementing PTT that way, promulgated by an organisation called OMA. However, the concept of doing this is a generic one. There was also at least one proprietary implementation of PTT over the mobile data network. It was called IDEN and came from Motorola. I will use the phrase "PTT over cellular" (lower case) to refer to the generic idea. Although the 2004 OMA protocol was a voice system, by the priority date the idea of incorporating video into the OMA protocol was being proposed publicly.
33. PTT over cellular was not successful in the UK at the priority date. Mr Unitt gave evidence of an attempt by a company called Dolphin which had limited customers (local authorities and a private ambulance service). I suspect (although it does not matter) those kinds of customer continued to use walkie talkie radios instead. However, PTT over cellular was in use in the USA and I accept Dr Kindberg's evidence that the skilled team based in the United Kingdom would know as a matter of common general knowledge what PTT over cellular was, and that voice protocols existed for it. A skilled team who had their attention drawn to "PTT" at the priority date would be interested in PTT over cellular. They would be able to and in fact would find the protocols, as well as the then current proposals for the future associated with the protocols, in order to implement it. If they did this the team would necessarily encounter the idea of doing video PTT. Their skills would allow them to implement such a system if they wished to do so. It would be a lot of work but that is a different matter.
34. There was a debate about whether the idea of using a PTT type system in an emergency or first responder type environment was common general knowledge. Although at times Mr Unitt seemed to express the view that that idea was not common general knowledge but only derived from the Voxer patent, I believe at day 1 p308 ln2-6 he accepted that that was, as a matter of common general knowledge, a classic implementation of this technology. In my judgment it was.

The patent

35. Paragraph [0001] describes the field of the invention as follows:
- “This invention pertains to telecommunications, and more particularly, to a telecommunication and multimedia management method and apparatus that enables users to review the messages of conversations in either a live mode or a time-shifted mode and to transition the conversation back and forth between the two modes, participate in multiple conversations and to archive the messages of conversations for later review or processing.”
36. The passage explains that the invention enables two modes of conversation. They are a live mode and a time-shifted mode. The user can transition between the two modes, participate in multiple conversations and archive the messages. The paragraph uses the word “review” to refer to what, in a voice system, would be the user listening to the words spoken in a particular message from which the conversation is composed. This word “review” would be understood as a generalisation of a user listening (to voice), watching (video) or reading (text). It is not being used to convey any sense of the timing of when the review has to take place. That is why the term “later review” is used in the last sentence of the paragraph.
37. There is then a description of related art which includes passages asserting that the current state of voice communication suffers from inertia ([0002]), refers to the drawbacks of existing voice mail systems ([0003]-0005), discusses current telephone systems ([0006]-[0009]), and then discusses “tactical” radio systems such as those used by the emergency services at [0010] –[0012]. The idea of using multiple channels is mentioned as is the lack of management tools that effectively prioritise messages ([0013]). Packet based networking and the concept of VoIP systems are mentioned. The problem which latency in a packet network running TCP causes for voice communications is referred to at [0014]. Further prior art is at [0015]-[0017].
38. The summary of the invention section has a number of consistory clauses and then a further paragraph [0022]. This emphasises the advantages which storage of the media created or received by the communication device provides. There is another mention of a variety of modes of conversations – live or time shifted. The invention is said to be applicable to phone calls, conference calls, instant voice messaging and tactical communications. It also provides for the ability to seamlessly transition back and forth between the two modes. The advantages of local storage on the communication device itself are referred to in the penultimate sentence.
39. The patent then turns to the specific embodiments and the drawings. The detail starts at paragraph [0026] with a section called “A. Functional Overview”. This section is akin to the functional specification of an IT system. According to paragraph [0026] the invention supports new modes of engaging in voice conversations and/or managing multiple simultaneous conversations. This can use a variety of media types including voice, video and text. Depending on their priorities recipients might participate in real time or be notified that a message is ready for later retrieval.
40. Paragraph [0027] is an important paragraph. It provides:

“[0027] Users are empowered to conduct communications in either: (i) a near-synchronous or "live" conversation, providing a user experience similar to a standard full duplex phone call; or (ii) in a series of back and forth time-delayed transmissions (i.e., time-shifted mode). Further, users engaged in a conversation can seamlessly transition from the live mode to the time-shifted mode and back again. This attribute also makes it possible for users to engage in multiple conversations, at the same time, by prioritizing and shifting between the two modes for each conversation. Two individuals using the system can therefore send recorded voice messages back and forth to each other and review the messages when convenient, or the messages can be sent at a rate where they essentially merge into a live, synchronous voice conversation. This new form of communication, for the purposes of the present application, is referred to as ‘Voxing’”

41. The first sentence in this passage is clearly talking about the same two modes of conversation which the previous passages have already referred to. Those two being live and time shifted. That matters because what had previously been referred to as a live mode now has the term “near-synchronous” applied to it – which conveys the idea that it is not necessarily fully “live” but perhaps somewhat stilted. The sentence also tells the reader something about the use of the inverted commas around the word live – both here and in earlier paragraph [0022]. The term live is being used in a figurative sense to encompass communication which the skilled reader might think a pedant would not have regarded as fully live.
42. The paragraph again refers to the idea of seamless transition and asserts that this feature allows users to engage in multiple conversations. It also characterises what is going on as the two individuals sending recorded voice messages back and forth between them. The messages can be reviewed when convenient or can be sent at a rate in which they merge into a live synchronous conversation.
43. The paragraph refers to the new form of communication as “voxing”, which is then explained in paragraph [0028] as a conversation consisting of a series of discrete recorded messages stored not only in the sender and receiver’s device but in the servers on multiple transmission hops across the network. The skilled reader would regard the language “voxing” with some suspicion and as hype. The idea of communicating by exchanging discrete recorded messages would not be seen as a new one. It is voice messaging, which the patent itself acknowledges is known. Therefore, whatever this new form of communication is, it must be more than that. The characteristic of most apparent importance in paragraph [0028] is that the messages are recorded (saved) in a number of locations. There are then further features addressed from line 42 of the paragraph. This states:

“Unlike a standard phone call or voice mail, the system provides the following features and advantages:

- (i) the conversation can transition between live and time-shifted or vice versa;

(ii) the discrete messages of the conversation are semantically threaded together and archived;

(iii) since the messages are recorded and are available for later retrieval, attention can be temporarily diverted from the conversation and then the conversation can be later reviewed when convenient;

(iv) the conversation can be paused for seconds, minutes, hours, or even days, and can be picked up again where left off;

(v) one can rejoin a conversation in progress and rapidly review missed messages and catch up to the current message (i.e., the live message);

(vi) no dedicated circuit is needed for the conversation to take place, as required with conventional phone calls; and

(vii) lastly, to initiate a conversation, one can simply begin transmitting to an individual or a group. If the person or persons on the other end notice that they are receiving a message, they have the option of reviewing and conducting a conversation in real time, or reviewing at a later time of their choice.”

[the separation into discrete sub-paragraphs has been added]

44. Feature (i) is the idea of a seamless transition between the two modes of conversation. That conversational context matters because seamless transitioning in the context of one-way broadcast live streaming, in other words the VCR functionality, would be regarded by the reader as common general knowledge.
45. Feature (ii) is semantic threading of the messages in a conversation.
46. Feature (iii) is recording for later retrieval. Features (iv) and (v) explain that as a result the conversation can be paused for as long or short a period as a user may wish, then picked up again and seamlessly transitioned to the live message.
47. Feature (vi) explains that no dedicated circuit is needed for the conversation as required by conventional phone calls and so, as feature (vii) explains, to initiate a conversation with an individual or group one can simply start transmitting.
48. Paragraph [0029] describes the concept of sending lower quality media quickly if network conditions are poor but then a higher fidelity “exact” copy later when network conditions allow.
49. Paragraph [0030] notes that the messages of conversations may be voice only or may include video and other data too. The reader would see that the patent takes for granted that a skilled person who was able to implement this system for voice would be able to implement it for video without any further help in the document. Another list of now familiar features appears in [0031] but nothing turns on it.

50. The next section of the patent is an extensive glossary. It starts with “Client”, which means the user’s application running in their device and then defines “Device” and “Server” in unsurprising ways. The term Message is defined as follows:

“**Message:** An individual unit of communication from one User to another. Each Message consists of some sort of Media, such as voice or video. Each Message is assigned certain attributes, including: (i) the User sending the message; (ii) the Conversation it belongs to; (iii) an optional or user created Importance Tag; (iv) a time stamp; and (v) the Media payload.”

51. Consistent with the focus of the patent on conversations, amongst the attributes of each message is which conversation it belongs to. The term “Conversation” is defined in a way consistent with that, as follows:

“**Conversation:** A thread of Messages (identified, persistently stored, grouped, and prioritized) between two or more Users on their Devices. Users generally participate in a Conversation using their Devices by either Reviewing Messages in real time or in a time-shifted mode, or creating and sending Messages of a Conversation as desired. When new Messages are created, they either define a new Conversation, or they are added to an existing Conversation.”

52. The two modes are referred to in slightly different terms here (the term “real time” used instead of “live”) but the two modes are obviously the same two as referred to elsewhere.

53. Skipping over numerous defined terms, the next one of significance in this case is Minimum Time Shift Delay (MTSD), which essentially means the delay inherent in the way it sends relevant data packets. Again skipping more definitions, an important one is Time Shifting (in paragraph [0038]) in which the glossary states:

“**Time Shifting:** Time shifting is the ability to play any Message at any time after it has been received as determined by the User-recipient. By Time-Shifting, a User may Review a Message: (i) immediately on demand by Rendering immediately after the MTSD; or (ii) time-shifted in a mode of reviewing the Message upon the discretion of the User; (iii) from the archive for searching, reconstructions, etc. of old Conversations; (iv) after a delayed period of time to accommodate the Reviewing of other higher Priority Messages (or Conversations) that need to be reviewed first; (v) and/or repeatedly if necessary for the Message to be reheard and understood. In other words, Time Shifting is the ability of a user to render a Message at any time after the system imposed MTSD”

54. In this definition the term “Rendering” just means playing. Clauses (ii), (iii), (iv) and (v) all make sense as examples of time shifting caused by the user but, at least at first sight clause (i) looks a bit odd as an example of what the user can do by time-shifting. However, what the patent would be understood to be trying to say is that by having a

system which can do time-shifting, all five things are possible. Whether the first one is really time shifting or not does not matter. In any event it sets the scene and provides the contrast with the other four so that they can be understood.

55. Facebook argues that in this definition time shifting is not caused by the network delay but is rather focussed on an extra shift in time determined by the user recipient. I accept that, looking at this passage. However, jumping ahead, there are later places in the specification in which time shifting is discussed in such a way that it includes the shift in time caused by network delay e.g. [0069]. The relevance of all this arises on construction of claim 1.
56. Before leaving the definition section I will say that I reject the suggestion by Facebook that there is some special legal principle of construction applicable to express definitions in patent specifications. There is not. The fact a term is given an express definition in the specification is obviously relevant when one comes to decide what the skilled reader would think the patentee meant by the words of the patent claim, and it could be determinative. However also relevant are all the other well-known things such as the specification as a whole and the common general knowledge.
57. After the glossary, from paragraph [0039] the specification describes the system and then the client architecture. An important aspect of the client architecture is the module called “Conversation/Message Management Services”. This module consists of a set of functions which manages the receipt and the sending of media ([0051] ln 33 and [0052] ln 50). Paragraph [0053] gives further information about this as follows:

“[0053] With the Conversation/Message management services 20f, all Conversations are essentially asynchronous. If two Users are actively engaged in a given Conversation and the User controlled delay between transmissions is minimal, the experience will be one of a synchronous full duplex conversation, as with current telephone or VoIP conversations. If either User delays their participation, for whatever reason, the Conversation drifts towards an asynchronous voice (or other Media) messaging experience. In alternative embodiments, Conversations can be optionally Tagged as asynchronous Messages only or synchronous Messages only. In either of these cases, the Conversation cannot drift between the two modes, unless the Tag is reset. After the Tag is reset, the Conversation again may flow between near synchronous (i.e. live or real-time) and asynchronous (i.e., time-shifted or voice messaging) modes.”
58. At first sight there is something confusing about this paragraph since it starts by saying that all conversations are essentially asynchronous but then later refers to the experience in some cases as being of a synchronous full duplex conversation. However, in the end the reader would understand what the patent is trying to say in the first two sentences clearly enough. The communication system does not demand an inherently synchronised relationship between the various messages in a conversation, for example a voice message from one speaker and the reply voice message from another. If in fact each is transmitted and received and chosen to be listened to quickly enough then the delays may well be so low that what the users actually experience feels like a true full

duplex conversation like a telephone call – as if they are in the same room simply talking to one another. So the conversation is essentially asynchronous even though the user experience is synchronous.

59. The third sentence of the paragraph then explains that if one user delays their participation then the experience will drift into being asynchronous, in the sense that a user will be aware of the delay.
60. The final two sentences refer to an alternative embodiment in which conversations are tagged so that they cannot shift between the two user experiences. In this context two kinds of messages are referred to – synchronous messages and asynchronous messages. The former are associated with the live real time mode and the latter with the time shifted mode or voice messaging. However, I do not believe this would be understood as ruling out the idea that it is possible to have a live or a “near synchronous” experience using asynchronous messages.
61. The next passage worth mentioning is [0063] which relates to the Messages/Signals Services module in the client architecture. This contemplates that there is the ability to signal the presence or absence of users on the network, to “ring” users to get their attention and to leave messages for users currently not on the network for them to review next time they connect.
62. Paragraph [0069] describes features of another module in the client architecture called Store and Stream. This mentions advantages provided by storage of media in the user’s device. In this passage there is a reference to “time shifted delivery” of a message as a result of lack of network connectivity. Voxel points out that this usage encompasses network delay within the concept of time shifting. Facebook says this should not be understood as use of the defined term, which relates to something a user does. Facebook is right in a grammatical sense but looking ahead to the dispute about claim construction, if the answer is to be found in that kind of meticulous verbal analysis (which I doubt) the language of the claim is concerned with a time shifted mode and a time shifted message rather than being concerned with a user engaging in the act of time shifting. There is another reference to a time shifted mode at the end of a later paragraph [0116] which confirms that the media can be retrieved from the local store or the one on the network.
63. At paragraph [0072] is a description of another important module in the client architecture called the “Persistent Infinite Message Buffer (PIMB)”. This is just an indexed data store with a fancy name. It is infinite in the sense of being unlimited in size and, if the device runs out of capacity, then data can be stored on the network. Data can be stored in a manner which is “arbitrarily persistent” meaning it is available “virtually forever”. In an alternative embodiment the data can be stored for a designated period of time e.g. 30 days. The specification also warns that the terms “persistent” and “infinite” should not be construed literally as absolute terms [0088].
64. As the specification explains in the context of the server architecture (from [0083]), there is a PIMB on the client device (PIMB 30) and a PIMB on a network server (PIMB 85).
65. Still in the context of server architecture, paragraph [0097] describes what is said to be a further unique aspect feature of the invention:

“[0097] One further unique aspect of the system 10 is that the media payloads generated by a Client 12 are stored in multiple locations. Not only are the payloads stored in the PIMB 30 of the generating Device 13, but also in the PIMB 85 of the Server(s) 16 and the PIMB 30 of the receiving Devices 13. This basic feature enables or makes possible much of the Voxing functionality described above and provides the system 10 with both resilience and operability, even when network conditions are poor or when a Participant of a Conversation is not connected to the network.”

66. Echoing paragraph [0028], this passage again refers to the importance of storing messages in multiple locations – in this case all the PIMBs – the ones in the two user devices and the ones in the servers.
67. The specification continues to paragraph [0215] but there is no need to examine those passages.

Claim construction and amendments

68. The law relating to claim construction is familiar and does not need to be set out here.
69. Although claim 1 of the re-amended claim set is in Annex A showing the amendments, it is useful to label the integers of the claim for the purpose of analysis as follows:
 - 1a A media communication method,
 - 1b which supports a live communication mode
 - 1c and at least one time-shifted communication mode,
 - 1d for communicating both voice and video media
 - 1e on a first communication device (13)
 - 1f over a communication network (14),
comprising
 - 1g progressively encoding, progressively and persistently storing on the first communication device (13) and progressively transmitting media of an outgoing message originated on the first communication device over the communication network, as the media is created; and
 - 1h progressively receiving, progressively and persistently storing on the first communication device (13) and progressively rendering media of an incoming message received over the communication network at the first communication device as the media is progressively received in a real-time rendering mode,

- 1i wherein the outgoing message and the incoming message are asynchronous messages (such that the media of an incoming message may be time-shifted with respect to the media of the outgoing message)
 - 1j that are transmitted over the communication network from the first communication device to the second communication device and received over the communication network at the first communication device from the second communication device without first establishing a connection over the communication network between the first communication device and the second communication device
 - 1k and wherein the outgoing message and the incoming message are stored and transmitted at each hop along a path over the communication network.
70. The introductory part of the claim provides that it relates to a media communication method (1a) and refers to a first communication device (1e). The communication will be between the first communication device and a second communication device (mentioned later in the claim at (1j)) over a communication network (1f). The method has to support a live communication mode (1b) and also at least one time-shifted communication mode (1c) and the method has to be for communicating “both” voice and video media (1d). The scope of all three of 1b, 1c and 1d are in dispute.

1a A media communication method

71. Voxer suggested, supported by Mr Unitt, that the invention would be understood to be a fundamentally new form of communication. That may or may not be right overall, but to the extent it is suggested that this has a bearing on the meaning of these first four words in claim 1, I do not agree. These words simply require there to be a media communication method as specified by the rest of the claim.

*1b live communication mode
and 1c at least one time-shifted communication mode*

72. The interpretation of these two features is best considered together. Starting with “live communication mode”, this would be understood as a reference to the live mode of having a conversation mentioned in the specification by contrast with the time shifted mode. It is the same mode whether it is called any of the following: live; “live”; real time; near synchronous or “live”; near synchronous “live”; near-synchronous real time “live mode”; a near synchronous, full duplex conversation (similar to standard “live” phone calls); near synchronous (i.e. live or real-time); or near real-time. All of these expressions appear in the specification.
73. It bears pointing out, if it matters, that the live communication mode is not limited only to a “nearly” live mode but would be understood as including a “truly” live mode in which two people were able to converse without being aware of any delay at all – i.e. like a traditional full duplex phone call.

74. How much delay is still within the ambit of a live communication mode? I will return to this below but at this stage it bears pointing out that there is a relationship between the answer to this question and the meaning of “time-shifted communication mode”. Facebook argues that time-shifted communication mode refers to the same thing as what they contend is the meaning of time-shifting in the glossary in the patent (para [0038]). The argument is that time-shifting does not mean just any shift in time between the utterance of a message and its rendering to be heard by the receiving user, it has a more limited meaning based on the glossary, of only a delay caused by a user choosing to render a message later than the moment in time it is received at that user’s device. In my judgment that is wrong for three reasons. First, the natural way the words themselves would be understood by the skilled reader would be as a reference to any kind of time shift. Second, the reader would not think the words in the claim were intended to limit a time shifted mode in the manner alleged when the patent specifically describes a communication mode in which the delay is caused by the network (paragraph [0069]). Third, the paragraph in the glossary makes sense as an explanation of what is possible without being seen as an attempt to put an artificial limit on the scope of what the claims mean by a time-shifted mode of communication.
75. I believe the distinction between the live mode and the time-shifted mode is simply a matter of the degree of time shifting. Nearly live will still be within the live mode but appreciable time shifting, whatever its cause, will take the conversation into the other mode. The claimed method requires the system to be capable of doing both. Therefore, if there is a significant delay from the point of view of the experience of the participants that will be the time shifted mode not the live mode.
76. On the other hand, if Facebook’s approach to time-shifted mode was right, that that mode is not characterised by the degree of time delay experienced, it would follow that the mode was instead characterised by user action. In which case the distinction between the live mode and time-shifted mode would not be one of degree but one of kind. That would be confusing and is another reason why I believe Facebook’s interpretation of time-shifting is wrong. It would be confusing because, if the circumstances worked out appropriately, a mode in which the user was making an active choice to listen to an incoming message as soon after being told it was received as possible, would be a time-shifted mode even though the experience was nearly live. Also, it could lead you to think that a message which had been delayed by a long time, say hours, but not due to any choice by the receiving user – might either have to be regarded as “live” even though no-one would ordinarily call it even “nearly” live, or would be a third mode of communication even though it is pretty clear the patent as a whole regards the two modes as covering a continuum.
77. It may be noted that the word “conversation” does not appear in claim 1. However, I believe Voxer accepted that the claimed invention is all about a method of carrying out conversations. For example, in opening, Voxer’s summary of what the Patentee’s new approach was said to be was as follows:
- “The Patentee teaches a new approach. Rather than thinking of a conversation as a single synchronous event, the Patentee recognises that a conversation can be considered as being made up of a series of asynchronous messages between pairs of participants. In this way, instead of two people (A and B) being in a single synchronous conversation, their conversation is

treated as being made up of asynchronous messages from A to B and B to A. From a technological perspective, this enables A to transmit a message to B at any time irrespective of whether B is listening, and vice-versa.”

[Voxer opening para 10]

78. Nevertheless, in case there is any doubt, in my judgment the two modes referred to in claim 1 are and would be understood to be modes of carrying on a conversation. The fact the word used is “communication” does not mean the patentee would be understood to be trying to encompass a mode of communication which was not a part of a conversation. Moreover, the fact that the later clauses in claim 1 refer to both an incoming message from a second device and an outgoing message to that same second device, supports the understanding that what is being described is a method of conversing.
79. I would also hold that in order to decide whether the modes supported by the communication method are live or time shifted modes, it is necessary to examine the timing relationship between the incoming and outgoing messages referred to in the claim.

Is a live broadcast the same thing as a live communication mode?

80. Before leaving claim feature 1b it is convenient to confront the construction question which arises on infringement. As explained below, Voxer contends that the system known as “Facebook live” or “Instagram live” has a live mode of communication even though there is a delay between the speaker (assuming it is a speaker) and the listener.
81. Voxer started by making the point that the skilled person would understand that what constitutes ‘live’ communication is context dependent. Thus, Voxer pointed out that the experts agreed that in a highly interactive communication context, such as on a telephone call, to ensure users have a good ‘live’ experience end-to-end delays should be kept to no more than a few hundred milliseconds. They did agree and I accept that.
82. Voxer then submits that in less interactive communication contexts (such as watching a live broadcast on radio or television) it is quite acceptable to have several seconds of end-to-end delay without disrupting the “live” experience for the listener or viewer. After all such delays are routinely included, for example, to provide a mechanism to edit out accidental expletives.
83. It is true that there are broadcasts referred to as live which include such delays. They are live from the point of view of the broadcaster because they are performing live as their performance is transmitted. Users who receive these broadcasts have an experience which is regarded as live and is referred to as such, not least because there is no faster way for them or anyone else on the network to experience the performance. This would be common general knowledge. However, it is not a context which the skilled reader would understand the patent to be talking about when it refers to a live communication mode. As I have already explained, the patent is focussed on a conversation between two (or more) users and the term live is used in that context.

84. Looking at the broadcast itself, there is only one stream, which can be regarded as a single continuous message. The fact that the different parts of that stream maintain the same time relationship that they had when they were uttered does not make them “synchronous” or “live”. In order to work out whether the period of delay inherent in a method which works this way is suitable for supporting a live communication mode, one has to ask whether it is suitable for supporting a live mode of having a conversation. To answer that one has to consider what would happen if the listener to the stream decided to reply to the “message” by sending back some kind of reply. Even if the reply was instantaneous it would still have the initial period of delay built into it.
85. If the replying user used the same broadcast technique, with its built-in delay, to send a reply back to the original broadcaster then there would be twice the delay period before a speaker heard any reply from their interlocutor.

Two modes and same media

86. Facebook also submitted that the two modes must be present as part of the same method of communication such that the same media must be reviewable live and on-demand, (i.e. time-shifted). On that basis the ability to receive a live communication and then review different content on-demand would not suffice. I agree.

at least one

87. The reference to “at least one” in feature 1c leads to one of Facebook’s added matter points and will be dealt with there.

Id for communicating both voice and video media

88. Facebook suggested that this term would be satisfied by a method capable of communicating voice alone at one stage and separately, on a different occasion, communicating video pictures (and no sound) on another occasion even if it was not capable of communicating voice and video together at the same time. I think the Comptroller may have been concerned that the claim could be read that way too and raised this as a clarity concern arising from the amendments.
89. Read in context, the claim would be satisfied by a method which conveys both images and voice (audio) at the same time. The claim will not be satisfied by a method which can only ever convey voice but never video. Similarly, it would also not be satisfied by a method which could only ever convey images but not voice (audio). What the claim does not require is that even if the method in question can convey both video and voice at the same time, it does not fall within the claim unless that method is also capable of conveying voice alone without video.
90. I do not believe it is necessary to do so but if it is necessary, then I would hold the claim does also cover the method proposed by Facebook, i.e. a method which can do voice on one occasion and video on another occasion but never the two together.

The remainder of the claim

91. The claim then has two subclauses defining outgoing and incoming messages respectively (1g and 1h) governed by the word comprising and then further terms which

qualify those definitions governed by two whereins (1i and 1j together and then 1k). The latter specify further characteristics of both the incoming and the outgoing messages and also requirements relating to how they operate.

1g progressively encoding, progressively and persistently storing on the first communication device (13) and progressively transmitting media of an outgoing message originated on the first communication device over the communication network, as the media is created;

92. Feature 1g provides that the media of an outgoing message must be encoded, stored on the first device and transmitted over the network “as the media is created”. In other words, all three of these things are to be done at the same time that the media itself is being created, e.g. in the case of voice, as the words are being spoken. One consequence of this will be that the transmission of the outgoing message is brisk enough to allow for the live communication mode to be possible.
93. The three things are required to be done “progressively” which indicates that the message is sent whilst the media is being created, rather than waiting for the entire message to be created before encoding, storing and transmission begins.
94. The storing has to take place “persistently”. The purpose of storing persistently is so that the data which has been stored can be retrieved later by the user. That might be shortly afterwards or a long time later. Mr Unitt thought the term “persistent storage” was a term of art and referred to any storage in a non-volatile medium such as flash memory or a hard drive (which remembers even when the power is switched off), by contrast with storage in volatile memory (RAM) which forgets when the power is turned off. I do not accept his evidence that it is a term of art at all, but even if it is, there is more to the point than this. I agree that to satisfy the claim one would need to store the message in a non-volatile medium such as flash, but while that is a necessary condition it is not a sufficient one. This claim language also requires that the message is kept in order that it can be retrieved later by the user. That does not mean forever, but there must be a degree of persistence. So merely putting the data in a buffer or other temporary store, even if that was in fact implemented using flash memory, would not satisfy claim 1.

1h progressively receiving, progressively and persistently storing on the first communication device (13) and progressively rendering media of an incoming message received over the communication network at the first communication device as the media is progressively received in a real-time rendering mode,

95. Feature 1h is similar to 1g but relates to an incoming message received at the first device from the network. It also provides for three things to be done “progressively” (i.e. as the message is received). They are receiving, storing and rendering of the media of the outgoing message. The storing also has to be done persistently, which has the same meaning and purpose as in the first clause. Amongst other things this act of storing allows for a time-shifted retrieval from the local device. The receiving, storing and rendering are done “as the media is progressively received in a real-time rendering mode”, in other words, there must be real-time presentation of the incoming voice/video stream. Like 1g, this means that the device has the ability to work briskly enough to support a live communication mode.

96. The natural way to read the claim is that the incoming and outgoing messages are part of the same conversation. Although it is not stated in terms, the claim does make clear that the two messages are passing between the same pair of devices (feature 1j), and that only really makes sense as a requirement if one is thinking about the two users having a conversation. All the same, as I have construed the claim above in relation to live and time-shifted modes of conversing, this point may not matter.

li wherein the outgoing message and the incoming message are asynchronous messages (such that the media of an incoming message may be time-shifted with respect to the media of the outgoing message)

97. Feature 1i provides that both the outgoing and incoming messages must be “asynchronous messages”. The proposed amendment seeks to add further words here in brackets which Voxer calls an explanatory rider and which Facebook opposes.

98. Voxer contended that the relationship which puzzled Facebook, between a live communication mode and asynchronous messages, is the very essence of the new form of communication invented by Voxer, which Facebook and Dr Kindberg have failed to appreciate. Voxer contends that the inventors’ contribution is founded on the realisation that conversations need not be thought of as having synchronicity between the participants. A better communication method can be achieved without trying to reproduce the ‘in the same room’ effect in which there is an immutable timing relationship (i.e. a synchronous relationship) governing the transmission of the various participants’ speech. Rather, the patent treats the conversation as a series of directional messages between pairs of participants which do not have any fixed timing relationship between them. From the perspective of a single participant, this allows the timing of their outgoing messages and the incoming messages to shift with respect to each other. The claim embodies this by requiring that the outgoing and incoming message are ‘asynchronous’.

99. I agree with some of this, but not all of it, as I shall explain. Contrary to Facebook’s case I do not agree that the fact the claim requires the messages to be asynchronous necessarily creates a fatal inconsistency between that and the requirement for a live communication mode. That is because, as Voxer contends and the explanatory rider in claim 1 makes clear, the term “asynchronous message” in the claim does not mean a message which is in fact shifted in time or “asynchronous” in that sense, it means a message whose timing can be varied so that the media of an incoming message can be (but need not be) time shifted relative to the outgoing message. This chimes with a passage in paragraph [0053] which provides that the conversation may flow between “near synchronous (i.e. live or real-time) and asynchronous (i.e. time-shifted or voice messaging modes)”.

100. If it so happens that the asynchronous messages are sent, arrive and are rendered fast enough so that the users’ experience of the conversation is as if they are in the same room, as in a standard telephone call, then the messages were nevertheless still asynchronous messages because they had the capacity to be slowed down or reviewed later. Paragraph [0053] does contemplate an alternative in which the messages are forced to be synchronous in nature so that the same room effect, as best it can be, is achieved all the time but that is not what the claim is focussed upon.

101. Whether this thinking is new or non-obvious is another matter.

102. It also bears making the point at this stage that in reaching these conclusions I have had in mind Facebook's clarity objections to the amendments. Facebook's own summary of its clarity objections to the amendments to claim 1 are as follows (written closing para 560):

'(a) If "asynchronous messages" have to be implemented in a "live communication mode" (rather than a time-shifted one) then it is unclear what this means.

(b) The proposed definition of "asynchronous messages" by reference to media which "may be time-shifted" suggests optionality, overlaps with synchronous messages, and lacks clarity.

103. I reject both points essentially for the reasons already explained above. Nor do I believe (if this was part of Facebook's case) that the skilled reader would find the claim particularly difficult to interpret in order to reach the relevant conclusions.

Ij that are transmitted over the communication network from the first communication device to the second communication device and received over the communication network at the first communication device from the second communication device without first establishing a connection over the communication network between the first communication device and the second communication device

104. This feature requires that the transmission and reception of the messages between the two devices occurs without first establishing a connection over the network between the two devices. Facebook's submission was, I think, that this means that the approach to communication must be a client-server model (in which each device is connected to a server or servers on the network) rather than based on a peer to peer model in which there is a direct connection between the two devices. However, asking questions about whether connections are "direct" tends to confuse the issue since that concept is not in the patent.

105. Voxer submitted that it is clear from the inventor's purpose and the claim language that what is required is that the first device must be able to send an outgoing message (i.e. one containing media) which is meant for the second device without establishing an end-to-end connection with that second device, i.e. regardless of whether the second device is even connected to the network at all. I agree.

106. As a consequence, it is hard to imagine how this could ever work without, in effect, a client-server model. That would allow the first device to interact with a server, after all the first device has to be able to send the outgoing message somewhere without first establishing a connection to the second device. However, that is different from saying that any system based on a client-server model will necessarily infringe. It will not. One could, for example, build a system on the client-server model which still required an end to end connection between the two user devices to be established first (via the server) before the first device was able to transmit at all, even to the server.

107. On the other hand, the familiar experience of sending SMS text messages on a mobile phone is, I believe, an example of what the patent is talking about. The text will be transmitted from the sender's phone without the sender first having to ring the

receiver's phone or the sender ever being concerned with whether the receiver's phone is switched on. The message is still transmitted from the sender's phone in such a case. The sender no longer has to worry about it and can turn off their phone. The text is stored somewhere on the network and the receiver will get the text when they switch on their phone. That would satisfy feature 1j.

108. Whether the first and second devices (when they are connected) are connected to the same server or different servers does not matter.
109. Of course, ultimately both devices will need network connectivity to transmit or receive messages at all, but that is another matter. Also, in the live communication mode, realistically both devices will be in operation, connected to the network, at the same time and exchanging messages. Facebook refer to that as a kind of indirect connection. I suppose it is, but I do not believe it is relevant. Feature 1j requires that the method must work in such a way that the messages can be transmitted without first establishing a connection between the devices.

1k and wherein the outgoing message and the incoming message are stored and transmitted at each hop along a path over the communication network.

110. Clause 1k provides that both messages must be stored and transmitted at each hop over the communication network. A "hop" on the communication path refers to a relevant server in the network. This language requires there to be storage (and transmission) at every one of the relevant servers in the path. It is necessary to add the word "relevant" because the skilled reader would know that telecommunications networks can involve many servers as part of the underlying network infrastructure, through which data will pass. The claim is not referring to those. The relevant servers are the application servers, i.e. they are the servers in the server architecture of the application which implements the communication system itself. In the specific embodiments in the patent specification the storage happens in the PIMB 85 on Servers 16 in the Server Architecture (see [0083] – [0088]).
111. The dispute on construction is whether an application server which does not have any storage, or putting it another way, does not have a PIMB, still counts as a relevant server (or a hop, which is the same thing). Voxer's case is that application servers without storage are not hops. Voxer supports this with the submission that Mr Unitt's evidence was that a hop in the patent means a "server with a capital S that includes storage". That was his evidence, but it is irrelevant. Hop is not a term of art. Voxer also submits that Dr Kindberg agreed in cross-examination that the hops in the specification included storage. I accept that, but it does not mean that the word hop in the claim – which is obviously a metaphor – would necessarily be understood as excluding application servers which did not contain storage.
112. In the end the question is – what would the skilled reader understand the patent to have used the words, requiring that the messages are stored and transmitted at each hop along a path, to mean? The objection to Voxer's definition is that it is circular. I do not agree. What it requires is storage at each application server which has the capacity to undertake that storage. If the server does not have that capacity, then it is not a hop. But if an implementer had application servers in the path which did have that capacity, but did not store the messages at those hops, then the set up would be outside the claim at least on a normal construction.

Claim 1 – clarity

113. I address the clarity objections to the amendments of other claims below. However, having completed the analysis of the construction of claim 1, it is convenient to record that I rejected Facebook's clarity objections to the amendments which put claim 1 into the form under consideration.

other construction issues

114. Neither side suggested that the construction of claims 2, 3 or 4 of the re-amended set had any bearing on the construction of claim 1 (or a later claim) and so there is no need to consider those.
115. Claim 5 requires the ability to seamlessly transition between the live and time-shifted communication modes. So far as I am aware there is no difficulty about this. There is a debate about whether this feature would be satisfied by a transition from a conversation happening in one mode to a different conversation in a different mode. I do not believe so. The issue is addressed in the section on added matter below.
116. Claim 10 requires queuing of the media on the communication device if the network connection is not available so that it is then transmitted progressively from persistent storage as soon as the network connection is available. Facebook was not alleged to infringe this claim and so did not attack its validity in this STS trial. Voxer maintains it is independently valid. So far as I am aware there is no difficulty about its construction.

Amendment

117. The amendments are objected to for added matter and lack of clarity.

Clarity of other claims

118. The law is that although lack of clarity in claim language *per se* is not a ground on which a claim can be found invalid, unless it amounts to a species of insufficiency (which species is now being referred to as uncertainty, for the sake of clarity), nevertheless lack of clarity in claim language which would be introduced by an amendment is a reason for refusing that amendment under s75 of the 1977 Patents Act.
119. Facebook takes a number of points on amended claims 2-5, addressed below.
120. The first point on claim 2 (Facebook's Amended Statement of Opposition para 13(a)) is that the requirement in claim 2 referring to the messages being near synchronous is either the same as the live communication mode of claim 1, and duplicative, or different from it and thereby creates ambiguity. I disagree. The skilled reader would understand that the live mode claimed in claim 1 has the meaning I have ascribed to it already. It includes a situation in which the users perceive no delay at all, in other words they perceive the same experience as a traditional full duplex telephone call. It also includes an experience in which delay is perceived albeit the delay is small and the experience is nearly live. No ambiguity is caused by claim 2 referred to the sub-set of the live communication mode which the patent calls near live in which the messages are near synchronous.

121. The second point on claim 2 (Facebook's Amended Statement of Opposition para 13(b)) is that the reference in claim 2 to the device using the real time rendering mode to do the progressive rendering required lacks clarity because it is duplicative of the similar language in claim 1. I think it probably is duplicative, but it does not justify refusing the amendment since, for the reasons expressed in the previous paragraph, claim 2 does contain a limitation with respect to claim 1.
122. The first point on claim 3 (Facebook's Amended Statement of Opposition para 15(a)) is that the requirement that the messages are near synchronous, even though the claim refers to the time-shifted mode (contrast claim 2) creates ambiguity. I disagree with this. As Voxer explains, what the reader would understand the claim to be referring to is the case in which the messages were received, stored and rendered "live" but are then re-reviewed from storage some period of time later in a time-shifted mode. Thus "near synchronous" in both claims 2 and 3 bears the same meaning and there is no ambiguity.
123. The second point on claim 3 (Facebook's Amended Statement of Opposition para 15(b)) is an added matter argument and will be addressed below.
124. The first point on claim 4 (Facebook's Amended Statement of Opposition para 17(a)) is the submission that the reference to messages being "not near synchronous" (as a result of being time shifted by storage on the network) introduces ambiguity because the difference between "not near synchronous" and asynchronous is not clear. I reject this. For one thing, as explained above, the reader would understand that in the specification (and claims) an asynchronous *message* is one which has the capacity to be time-shifted. In that context (which is how it appears in claim 1) the word "asynchronous" is not being used to refer to the actual timing of a given message. Whereas claim 4 is talking about the timing of an asynchronous message which has in fact been time-shifted in a particular way (due to network storage) with the result that it is not near synchronous at the first communication device. Furthermore, as Voxer contends, some asynchronous messages may be in fact be close to synchronous (i.e. near synchronous) and other asynchronous messages may be further away from that (i.e. not near synchronous). I reject this objection.
125. The second point on claim 4 (Facebook's Amended Statement of Opposition para 17(b)) was put on a premise about a submission by Voxer about the meaning of claim 4 which Voxer expressly disavowed (Voxer opening paragraph 156) and so I do not need to examine it.
126. Facebook takes one clarity point on claim 5, contending (Facebook's Amended Statement of Opposition para 19(d)) that "to the extent that Voxer will place reliance on the 'seamless' nature of the transitioning between the two modes, such a requirement is not enabled, alternatively is conceptually uncertain and the specification of the Patent fails to provide sufficient explanation or direction to resolve the said uncertainty". The point being made here is a clear one in the sense that Facebook can imagine that Voxer might point to some transitioning between the two modes, say in a bit of prior art, and assert for some reason Facebook cannot predict that although it is a transition between the modes, for some reason it is not a seamless one. Then Facebook would have the ability to say that such a narrow meaning is not clear. However, there is no such point in the case and so this issue does not arise. In its written submissions (para 560 (d)) Facebook pretended that it was taking a wider and unconditional point that "seamless" was just ambiguous. In effect this was a repackaging of concerns raised by the

Comptroller which I will address below. Facebook did not plead this and it is not open to Facebook.

127. The Comptroller raised a number of objections, which Facebook summarised in its written closing at para 559 as set out below. They include added matter points as well as clarity objections but it is convenient to set them all out in one go:

“(a) Claim 1 requires the method to “support” live and time-shifted modes but the requirement for “support” in a method claim is conceptually unclear;

(b) The steps in claim 1 (“as the media is created...”, “in a real-time rendering mode”) are directed towards a live mode but it is unclear how the claimed steps support a time-shifted mode at all; it is therefore unclear how the steps of the method support both modes of operation, and adds matter;

(c) The parenthetical definition of “asynchronous messages” in claim 1 is unclear and “suggests that time shifting is somehow optional, which is contradictory”, while if this is meant to be a technical feature it is unclear;

(d) [*this related to Voxer’s first amendment application and no longer arises*]

(e) Similar objections based on lack of clarity [*to those raised by Facebook*] were made to proposed claims 2, 3 and 4. Claim 3 also added matter since it implies that claim 1 encompasses both rendering media “out of storage” when time-shifting and also something else (which, as noted above, must add matter).

(f) Proposed claim 5 adds the feature of “seamlessly transition[ing]” but this is “not a sufficiently enabling technical feature”, uses a subjective criterion which is not clearly defined, and lacks clarity. The amendment also adds matter by implying that users can transition between modes between multiple conversations, when the specification only discloses transitioning between modes within a conversation (with a separate step of switching to other conversations).”

128. Point (a) is about the word “support” in claim 1. Notably Facebook, which has not been shy of taking every conceivable point in this case and of adopting points made by the Comptroller, did not adopt this one. I am not surprised. Although I can quite see that in the negotiations which would go on between the Patent Office and an applicant, the Office might well try to push the applicant into using more precise terminology, in its context in this case the word is clear enough. It means that the claimed method must be able to allow each of those modes to take place.
129. Point (b) is a similar issue. The fact that the amendment to the claim requires the method to support two modes of communication, even though the remainder of the

claim focusses on steps relating to one of them, does not introduce a fatal ambiguity nor does it amount to added matter.

130. Point (c) is a different way of making the same point which Facebook advanced about the alleged inconsistency between asynchronous messages and a live communication mode. I have addressed and rejected this submission above.
131. Point (d) no longer arises and, on point (e), all the clarity points taken in relation to claims 2, 3 and 4 have been addressed already. A separate point on added matter is taken on claim 3 and is addressed below.
132. Point (f) is a set of arguments about “seamless” in claim 5. In terms of clarity, there is no difficulty with this when the common general knowledge of the skilled person is taken into account. The ability to seamlessly transition back and forth between live and time-shifted in the context of a single video stream is well known and understood as part of VCR functionality. The ability to pause live TV is an example. The seamless quality is from the perspective of the user but the system’s capacity to provide this is a technical feature. It is not subjective.

Added matter

133. The law on added matter is not in dispute. The test requires a comparison between the application as filed and the patent as sought to be amended (as a whole). The fundamental question is whether or not the skilled person would learn anything different from the patent as amended compared to what they would learn from the application as filed. Aldous J (as he then was) explained how to carry out the comparison in ***Bonzel v. Intervention*** [1991] RPC 553. The comparison is strict in the sense that subject matter will be added unless such matter is clearly and unambiguously disclosed in the application either explicitly or implicitly. This is the universal test for added matter. It applies regardless of whether the objection raised is said to belong to some particular sub-species, such as intermediate generalisation. The most up to date summary of the law in this area, picking up the other relevant cases, is in paragraphs 55-60 of the judgment of Floyd LJ in ***Conversant Wireless Licensing Sarl v Huawei Technologies Co Ltd*** [2020] EWCA Civ 1292.
134. Facebook submitted in its opening skeleton (para 293) that “the essential policy behind the rule against adding matter is to prevent the patentee from obtaining a different monopoly to that which the application originally justified, especially where that would not be ascertainable to third parties from the application”, citing ***Conversant*** at paragraph 55 in support (in fact the words relied on are based on para 10 of ***AP Racing v. Alcon Components*** [2014] EWCA Civ 40). To extract those words from the authorities and then place emphasis on them in the way Facebook has, risks making a mistake. There is no doubt that the law of added matter is there to protect third parties by limiting the patentee’s room for manoeuvre by reference to the application as filed. It is also true that it prevents the patentee from obtaining certain kinds of different monopolies. But some differences between the monopoly claimed in the granted patent as opposed to whatever was claimed or disclosed in the application as filed are unproblematic and lawful, and are not prevented by the law of added matter. It is an error to focus simply on the idea of a difference in the monopolies and it is not what the previous cases were talking about. There is no substitute for applying the law as a whole as explained in ***Conversant*** using the comparison explained in ***Bonzel***.

135. The main objection to the 20th November 2020 amendments to claim 1 related to the phrase “*wherein the outgoing message and the incoming message are stored and transmitted at least one hop along a path over the communication network*”. Facebook contended that the application as filed only disclosed the idea of storing the message at every hop along the communication path and so the term “at least one hop” discloses the idea that one need not store at every relevant server and so is added matter. The Comptroller supported this objection. Sensibly Voxer fixed the problem in the re-amended claim set of 25th February 2021, which uses the term “at each hop” instead.
136. The remaining added matter objections to the amendments to claim 1 relate to:
- i) “*which supports ... at least one time shifted communication mode...*”
 - ii) “*asynchronous messages (such that the media of an incoming message may be time-shifted with respect to the media of the outgoing message)*”
 - iii) “*wherein the outgoing message and the incoming message are stored and transmitted at each hop along a path over the communication network*”
137. There are also objections to claim 3 as a whole and to the following language in new claim 5:
- i) “*seamlessly transition between the live communication mode and the time-shifted communication mode*”
138. I will take them in turn.

Claim 1: “*which supports ... at least one time shifted communication mode...*”

139. Facebook contends that this adds matter because the application as filed only discloses a single time shifting mode whereas this discloses the idea that there is more than one. Facebook points out that paragraphs [0001] and [0017] as well as claim 1 of the application as filed all refer to “a” time shifting mode, so for example paragraph [0001] talks about the invention pertaining to a telecommunication and multimedia management method that enables users to review the messages of conversations “in either a live mode or a time-shifted mode”.
140. I do not accept this. Facebook’s real complaint is that Voxer asserts that the invention is not limited to what Facebook regards as the originally disclosed time-shifted mode but rather includes different and additional “time-shifted modes”, such as a mode where the media is not first stored on the first communication device and is then later rendered from that storage, but rather is rendered immediately on receipt from the network. The construction of time-shifting which Facebook contended for and which I have rejected above would have meant that it excluded rendering a message immediately on receipt from the network and so Facebook’s added matter argument is presented as if it is related to its case on construction. In fact I believe this added matter argument is wrong irrespective of the construction issue. The glossary paragraph which Facebook relies on to support its case on the meaning of time-shifting (which is also in the application as filed at paragraph [0073]) positively describes a number of different things a user can do which are versions of time shifting. That shows that whatever time-shifting means, to refer to “at least one time shifted mode” is not added matter.

Claim 1: “*asynchronous messages (such that the media of an incoming message may be time-shifted with respect to the media of the outgoing message)*”

141. In its closing skeleton Facebook took three points. The first (para 501) and third (para 505) can be dealt with briefly. The first point was that “asynchronous messages” were not in claim 1 of the application as filed, but that is not a valid added matter objection. The third point was premised on “asynchronous message” equating to a message which is time-shifted by the choice of the receiving user, however that is not what I have held “asynchronous message” means.
142. The second point (para 502-504) relates to the explanatory rider in brackets, which is proposed to be inserted by amendment. It is the text which confirms that “asynchronous message” refers to a capacity of the message and not to whether the media of the message has in fact been time-shifted. This idea is said not to be disclosed in the application as filed. I disagree. In my judgment it is disclosed in the paragraph which is [0053] of the granted patent and [0095] of the application as filed. This paragraph makes clear that even if the user experience of a conversation is that of a synchronous full duplex conversation, the messages which are used to convey the media will be what the patent calls asynchronous messages. It also makes clear that when these asynchronous messages are used in a conversation the user’s choice can cause the experience to change from that of a synchronous full duplex conversation to a time shifted one. The messages in both cases are always “asynchronous messages”. In other words what the patent application calls “asynchronous messages” have the capacity such that the media can be time shifted or not. Thus, the words of the rider in brackets do not add matter. They are supported by the application as filed. The reader would not learn new information from claim 1 as proposed to be amended.

Claim 1: “*wherein the outgoing message and the incoming message are stored and transmitted at each hop along a path over the communication network*”

143. I think Facebook’s first point, which was something to do with claim 9 as granted, was that claim 9 as granted required both incoming and outgoing messages to be stored at each hop whereas, if these words did not do that there was an extension to claim scope or (maybe) an intermediate generalisation. It was said to be a squeeze on infringement but, while not impossible, squeezes between infringement and added matter are rare because claim scope is not the same thing as what is disclosed by a claim (see AC Edwards v Acme and Alcon). Moreover, comparisons with a granted claim are not relevant to added matter because what matters is the application as filed. In this case no objection of extension of claim scope (Art 123(3) EPC) has been raised. I reject the added matter argument because I can discern no material difference in disclosure between this proposed amendment and paragraph [0151] of the application as filed, which is the relevant paragraph.
144. Facebook’s second objection is that the text does not include the word “persistently” and so teaches the idea that the storage need not be persistent at each hop, whereas what is disclosed in the application as filed is persistent storage at each hop. I reject this because the skilled reader would understand this feature of the claim to be referring to the same persistent storage as is described at length in the application and is also referred to in feature 1g and 1h. Therefore, there is no added matter.

Claim 3 added matter points

145. Facebook's added matter point (Amended Statement of Opposition para 15(b)) is that the only support for new claim 3 to be found in the application as filed (as opposed to the granted patent) is claim 10 as filed and that this does not support the outgoing and incoming messages as near synchronous. Maybe it does not but I reject the argument that this idea in claim 3 amounts to added matter. What claim 3 discloses is the idea that the method supports a time-shifted mode in which the first device plays the media of the incoming message out of storage on that device after the incoming message was received (e.g. because the user has chosen to do this) in a case in which the outgoing message and the incoming message are near synchronous at the first communication device. This idea is clearly part of the general disclosure of the invention in the application as filed.
146. The Comptroller's concern is that if progressively rendering media out of storage (in claim 3) is in addition to rendering it in real-time as required by claim 1 then this seemingly adds matter and is unclear. I do not believe it is unclear, nor does it add matter. The method disclosed in the application is capable of doing both things. It must be able to render in real-time (claim 1) and it must also be able to render media from storage in a time-shifted mode (claim 3). Claim 3 does not contradict claim 1.

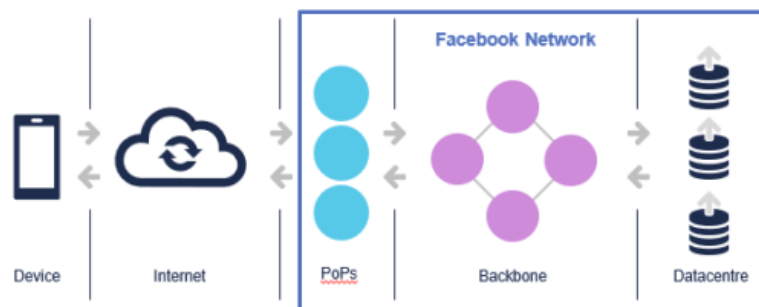
Claim 5: "seamlessly transition between the live communication mode and the time-shifted communication mode"

147. The added matter here is alleged to be that the transition is not limited to a transition taking place in a conversation. As I have construed the claim, this point does not arise because it is clear that the live communication mode and the time-shifted communication mode are modes of having a conversation. Thus claim 5 is claiming the ability to seamlessly transition in that conversation. The idea of the transition being seamless only makes sense when the live mode and time-shifted mode are in the same conversation.
148. A further point relates to transitions between conversations. There was a question whether transitioning between two conversations did or could result in a change of mode. Voxer pointed out that Figure 13C of the patent (and in the application as filed) does show a switch between conversations which also changes the mode from live to time-shifted. I agree that that is in Figure 13C but I do not agree that that is what claim 5 would be understood to be referring to. A method where the first communication device could not seamlessly transition between the live mode and time-shifted mode of the same conversation, would not be within claim 5. The claim is not concerned with transitions between separate conversations and debates about what is disclosed in that regard are irrelevant.

Infringement

149. To recap, the infringement case is advanced in relation to the Facebook website and the Facebook and Instagram Apps as they operate on iOS devices. In each case what is alleged to infringe is a live broadcast feature. The three live broadcast features are not identical to one another but they are very similar.

150. As is well known, Facebook and Instagram are social networks. That means a user who has an account on one of those networks is able to interact with other users on the same network. From the point of view of a user, put in broad terms there are two relevant classes of other people. The first is users with a specific relationship within the network. On Facebook they are called “friends” and on Instagram “followers”. The second is the public, i.e. anyone. Generally, and subject to privacy settings, the public includes not only other network users but other members of the public on the internet generally.
151. A user who is a friend/follower of another user may be automatically notified of any posts made by the first user. Examples of what can be posted include text, photographs and pre-recorded video clips. The networks have privacy settings but they are irrelevant to this case.
152. At a technical level the way these social networks function using the Facebook and Instagram Apps is that the App runs as application software on a user’s mobile device. A user with a Facebook account can also access it via a web browser. This can take place via a mobile device or a personal computer but using a web browser the live broadcast feature is only available from a personal computer.
153. A diagram illustrating the Facebook global network architecture is as follows:



154. “PoP” stands for Point of Presence. These are local servers operated by Facebook. They are local in the sense that the individual PoP which a user’s device will connect to will be one geographically close to the user’s device. Some requests from a device will be dealt with by the PoP but if not, the request will be relayed to the data centres via the backbone routing system.
155. The Instagram social network runs on equivalent network architecture.

The live broadcast feature

156. The live broadcast feature allows a user to broadcast a video stream to others. This is different from posting a pre-recorded video clip. Using this feature a live video is streamed from the user’s device. The stream is captured from the device’s camera and microphone. It is sent from the device to a server in the Facebook network called the Facebook Live Server (FBLs). Despite the name this applies to Instagram too. While the stream is being captured and sent, a limited number of video frames are temporarily

cached in a short-term buffer (in RAM). In addition, a complete copy of the video stream encoded in a different way is stored on the local device in a high-quality video format in a temporary folder. On a mobile device that data is stored in flash memory in an operating system folder called *tmp*. Video data is not stored permanently in this local temporary folder.

157. While the video is being broadcast, the broadcasting user cannot access any other functions in the Facebook or Instagram software.
158. The broadcasting user has some choice about the persons to whom the live video stream is made available (such as friends/followers only or to the public). The detail of this differs between Facebook and Instagram but the differences do not matter.
159. When the broadcast is finished the broadcasting user can choose to “delete”, “share” or “save” the video data. Deleting the video deletes the data from the temporary folder on the device. Sharing the video makes the video available on the user’s timeline. In effect it amounts to a post of the video. The video data is stored in the Facebook network and a link to it is placed on the user’s timeline as a “was live” video. Saving the video means the user keeps their own copy of the video data.
160. There are various ways in which another user of the social network may be alerted to a live broadcast. They do not matter. When another user wishes to do so they can click on a link to the live broadcast. The video data will then be sent from the Facebook/Instagram network to the viewer’s device where it will be decoded and rendered so that it can be viewed.
161. For various reasons there is a transmission delay of about 10 seconds between the broadcaster broadcasting the video and the viewer viewing it. The viewer can choose to pause, rewind or fast forward the video up to what is called the live head of the broadcast. The live head will always be at least that delay period later in time than the broadcast itself. The delay is inevitable and neither the broadcasting user nor the viewer can reduce it below 10 seconds.
162. The viewer’s device has a temporary caching system which stores a few seconds of the video stream being watched. The viewer can select a “save video” option which will allow the user to access the video afterwards if the broadcaster has chosen to share it but will not create a saved copy if the broadcaster does not do that. There is also a pre-fetch system which stores the first seconds of videos the system thinks the viewer may be going to watch, e.g. as they scroll through a timeline.
163. Assuming the broadcaster has chosen to share a copy of the broadcast on the social media network, then a viewer can choose to view that video by selecting it. They can again pause, rewind and fast forward that video. If the broadcaster did not choose to share a copy of the broadcast, then a viewer cannot replay the video even if some of the video data had been stored on their own device.
164. There is much more detail to how this all works but the above description is sufficient to understand almost all of the infringement issues. I do not believe anything above is justifiably confidential.

165. Before moving on to infringement, it is worth labouring one point. So far none of what has been described above mentions two people having a conversation. That is not an accident. In order for two people to have a conversation using this technology, the first user would have to broadcast a video stream I will call video stream A. The second user could access video stream A as soon as it was available or choose to watch it at a later time. However, they could not send a video back again until they stopped watching video stream A. Once they chose to do so, the second user could then start broadcasting their own video stream, call that video stream B. The first user could then, if they wished to do so, access video stream B as soon as it was available or choose to watch it at a later time. And the process could continue. That would be a form of conversation between the two users because of the semantic connection in the minds of the two users between the two video clips. Nevertheless, there is no link between these two video streams A and B at a technical level. This conversation is possible simply because everyone is able to broadcast a live video stream from their device in this system. Notably in this conversation there would be a minimum of 20 second round trip delay from the point of view of the first user and their viewing/hearing any reply from the second user.
166. A paradigm use for the live broadcast feature would be a musician playing music and broadcasting it to their fans on social media. Voxer repeatedly referred to the idea that in this context the fans would be able to interact with the musician while video stream A was playing, by using other functions in the social media network e.g. by posting a text request for another song to be played or by “liking” the video. That is so. It also amounts to a conversation of sorts and in that example the round-trip delay from the point of view of the first user is 10 seconds. Its relevance will be considered below.

The issues on infringement

167. The points to consider are:
- i) live communication mode
 - ii) time-shifted communication mode
 - iii) both voice and video
 - iv) persistent storage, outgoing message
 - v) persistent storage, incoming message
 - vi) incoming and outgoing messages
 - vii) asynchronous message
 - viii) stored at each hop
168. In addition to infringement on a normal construction under s60(1), there are two equivalent issues, issues on s60(2) and *Formstein*.

live communication mode

169. The issue here is the inevitable 10 second delay. The expert evidence does not assist on this. None of the relevant terms are terms of art. Facebook contends that this sort of delay is not what the patent means by live, particularly when one bears in mind that in a conversation between two users there will be a 20 second minimum round-trip time. Voxer's retort makes the reasonable point that Facebook itself uses the word "live" to characterise the feature alleged to infringe, however as I have already explained in the construction section above, the word live used in that broadcast context is referring to a different idea.
170. I find that the 10 second transmission delay inherent in live broadcast feature is sufficiently large that the methods alleged to infringe do not support a live communication mode as required by claim 1. The reason is because one needs to ask whether this system would provide a live mode for having a conversation, and the answer is no. The experience of a minimum of 10 seconds between whatever it is the broadcasting user wishes to communicate in their outgoing message and its receipt by the viewing user is too great a delay. That is so even if one imagines the receiving user somehow replying instantaneously, i.e. Voxer's example of a fan posting back a text request for the next track without using the live broadcast feature.
171. For what it is worth I would also hold that the right way to look at this is to examine the timing of a conversation in which both parties are communicating using the relevant method – in which case the round trip delay is 20 seconds. That makes it even clearer that this is not a method which supports a live mode for having a conversation.
172. Putting it another way, Voxer emphasised (closing para 65) that "the patent recognises that there will be some delay, and there is no warrant for any specific limit to be written into the claim – it all depends on context." I agree the claim does not contain a bright line limit measured in seconds, and I agree that the patent recognises that the inevitable delays inherent in such a network may lead to a delay discernible to the user. The line drawn by the patent is whether the system supports a live communication mode as the skilled reader would understand the patentee to have used those words to mean. A minimum 10 second delay is just too big.
173. Related to this was a point made by Facebook that the alleged infringing method was not "a communication method" at all because it was not a method of interactive communication. I have rejected that construction of the term above, but I will also say that I accept Voxer's example which I have explained with video streams A and B as showing that the system can support a method of interactive communication. The problem for Voxer is that that interactive communication does not have a live mode.
174. Voxer also argued that Facebook used the term live to apply to the live broadcast feature in order to emphasise the distinction between that approach and the existing one whereby pre-recorded videos are posted. Irrespective of Facebook's intentions, I agree that the distinction exists and that calling it a "live broadcast" serves to emphasise the difference. However, the fact this difference exists does not necessarily mean the Facebook feature has a live mode within the meaning of the claim.
175. Voxer also made a related point, emphasising the progressive nature of the transmission in the live broadcast feature and how that differs from a method involving posting pre-recorded video clips. The Facebook system is "progressive" in that sense referred to in the claim because it does start transmitting more or less instantaneously and continues

while the broadcasting user continues to speak, sing or whatever. However, the fact that kind of progressive transmission would be necessary for a live communication mode does not mean it is sufficient to satisfy the requirement of a live mode. It is not. Nor is it right (as Voxer alleged in paragraph 63(b) of its written closing) that it is relevant that the viewer is able to participate whilst the message is still being sent. That fact is an indication that the transmission is progressive but it does not establish that the method necessarily supports a live communication mode.

176. Voxer also refers to the ability in the Facebook system of the broadcasting user to invite another user to join the broadcast. This is called “live with”. Contrary to Voxer’s case, it does not show that the live broadcast is a live communication mode as claimed because it makes no difference to the transmission delay already referred to. Voxer rightly did not suggest this “live with” feature itself was a live communication mode.

time-shifted communication mode

177. The live broadcast feature clearly has at least one time-shifted communication mode. In fact, it has two, because the 10 second delay means that as I have construed the claim, it is a time-shifted communication mode. The fact the time shift is caused by the network and is not under either user’s control is irrelevant. There are also then other time-shifted communication modes, because users can pause the stream and then start watching it again (introducing more delay) and users can also watch the broadcast much later if the broadcaster has chosen to make it available that way.
178. Facebook sought to make a complicated point about the fact that the stored “was live” broadcast could be different from the actual live data stream. An example of why this might be so was because network quality might mean that the live stream a user received had had to be encoded at a lower bit rate (poorer quality) to accommodate that poor network connection whereas the stored “was live” video was high quality. That makes no difference in my judgment because what matters is that the media – i.e. what the user watches – is the same media from the point of view of the user. The fact the media may have been encoded in a different format, whether higher or lower quality or for some other reason, is irrelevant. The fact there may be minor differences between the way the media appears to users on different occasions also makes no difference.

both voice and video

179. I think Facebook had a point that this claim feature 1d meant that the system must be capable of supporting a voice alone method of communication, which the live broadcast feature cannot do. I agree the live broadcast feature does not have that capacity, but the claim is not limited in the way Facebook contends. The claim is satisfied by a method which only supports combined audio/video messages.

persistent storage, outgoing message

180. Facebook contended that there is no persistent storage of the outgoing message on the broadcasting user’s device. The only storage of outgoing data is either (a) storage in the short-term buffer in RAM (b) the storage of a temporary copy in the *tmp* directory of iOS devices, which is in non-volatile flash memory. Voxer accepts that storage in the short-term RAM buffer does not satisfy the claim. I agree. Voxer relies on the temporary copy in the *tmp* directory. Facebook says it is not persistent because Apple’s

guidance information shows that the *tmp* directory is periodically purged by the operating system when the app is closed or inactive, and when that happens the data will be lost. Moreover, says Facebook, that data is not accessible to a user. Voxer contends that this does amount to persistent storage because while an app is in operation the operating system will not purge the *tmp* directory.

181. My findings on this are as follows. The fact that the user cannot directly access the *tmp* data is not the whole story. The fact it is being kept while the app is running is no accident. One reason is in order to allow the broadcasting user to upload (share) a copy of the broadcast to the Facebook network in order to allow viewers to watch it if they wish. Another reason is to allow the broadcasting user to “save” the video at the end of the broadcast. As partly explained already, if the user chooses either of these options then the video file in the *tmp* folder will either be uploaded to the network for others to watch (in the first case) or be moved into local non-volatile storage outside the app (if “save” is selected) so that it can be accessed in the same way as other photos or videos saved on a mobile device.
182. In my judgment the act of storage in the *tmp* folder amounts to persistently storing the outgoing message as required by claim 1. It is not merely putting the data in a temporary store, it is doing so to achieve the purpose of allowing the message data to be retrieved later. The fact the user may not realise that is happening is irrelevant, so also is the fact one can conceive of some circumstances in which the *tmp* file could be lost. Looking at it the other way round, in a case in which the broadcasting user has chosen to upload (share) the video message to the network (and unbeknownst to the user that file came from the *tmp* folder), the reason that was possible is because the message had been progressively and persistently stored starting in the *tmp* folder.
183. There is no equivalent local storage using the website version of the live broadcast feature because the buffering and caching on a personal computer is all in RAM. Voxer relied on the fact that some data stored in RAM would end up on disk because computers have a system of swapping out data from RAM to disk if the RAM is full. The skilled person would not regard that as satisfying the requirement for persistently storing the outgoing data. Therefore, the website version of the live broadcast feature run on a personal computer does not satisfy this claim feature.

persistent storage, incoming message

184. Facebook contended that there is no persistent storage of the incoming message on the viewing user’s device. The only storage of incoming data is either (a) storage in the short-term look ahead buffer which is discarded when the live head is reached, (b) the storage in a temporary cache. It is common ground (as I understand it) that (a) is irrelevant. The issue is (b).
185. There are algorithms which limit the amount of storage devoted to a video in the temporary cache but the details of how they work and how the Instagram algorithm differs from the Facebook algorithm are confidential and irrelevant.
186. If a viewing user wanted to watch the video after it had finished, their device will access the “was live” video stored on the network. However the fact that it is Facebook’s network system which makes retrieval of the video media after it has finished possible

does not in and of itself show whether the local storage of the incoming video stream in the temporary cache is persistent or not.

187. Although Facebook sought to minimise it, in fact there are circumstances in which the persistence of the stored video stream in the cache matters. The cross-examination established that the following scenario occurs and is realistic. A viewing user can be watching a live video. As they do, the incoming stream is stored in the temporary cache. The viewing user can pause and rewind. They can start watching the stream again without fast forwarding back to the live head. In some cases, e.g. if the pause is for an appreciable time, then when the viewing user starts watching again what they will view will have come from the network. However, there are circumstances in which, when the viewing user starts watching again after a pause or rewind, then the stream they see will be rendered from data in the temporary cache. Thus, the storage in the temporary cache allows for later retrieval by the user, even if is only a relatively small period of time later from the user's perspective. That is enough to amount to persistently storing the incoming data. The fact it is later removed outside the user's control is irrelevant.
188. Facebook suggested this circumstance "verged on de minimis" but there was no evidence for that and I reject it. The time to take a non-infringement point based on de minimis was this trial. This capacity is a capacity provided by the method in issue. It satisfies this aspect of the claim.
189. Voxer perceived that Facebook took a further point that because time-shifting from network storage was possible as well, the claim could not be satisfied. One of Voxer's two equivalents arguments was premised on that point being accepted. I do not accept it. As explained above, the methods alleged to infringe have the relevant capacity. The fact that storage is not forever does not matter. It is persistent enough to provide a retrieval function. The fact that another retrieval function also exists based on network storage does not take the method outside the claim.

incoming and outgoing messages

190. The point here is that it is not possible in the Facebook system for the one device to send an outgoing message at the same time as it receives an incoming message. It is true that the system cannot do that, but I do not believe the claim requires it to be possible. In other words a full duplex approach would not, I think, be possible using Facebook's live broadcast feature. However what would in effect amount to a half duplex live communication mode, which would satisfy the claim, could still have been produced if the timing had worked out appropriately (in other words if the inherent 10 second delay had been significantly shorter).

asynchronous message

191. Before trial, Voxer's primary infringement case had not focussed on the ability to carry out a conversation using the Facebook system which I have described above. That "use case" only emerged at trial. In Voxer's primary infringement case before trial the outgoing and incoming messages were entirely unrelated. The point was simply that a mobile device running the app (or a personal computer logged into the website) could be used to send an outgoing message as called for by the claim and also, on another unrelated occasion, the device could be used to view a video broadcast by someone else. Facebook took a point on this integer as a vehicle for submitting that in Voxer's

primary infringement use case there was no semantic connection between the incoming and outgoing messages. Without that, it was said, they cannot be asynchronous messages because they have no relationship at all.

192. At first sight that point is wrong because as I have held already an asynchronous message is one with the capacity to be time-shifted and so, since the viewer is able to pause rewind and restart the video stream, the timing can be varied. So far so simple. However, as the rider in claim feature 1i makes clear (and was the case anyway from the specification – which is why the rider is not added matter), the timing shift in question which has to be possible in order to give the incoming message the relevant characteristic is a shift relative to the outgoing message. It is not talking about a shift in timing inside a video stream within a single message, relative to itself.
193. The difficulty arises because Voxer is trying to read the claim concerned with a system in which the communication between two users is actually a conversation broken up into a collection of discrete messages which do relate to one another, onto a video streaming system in which there is only a single message, which is the entire video stream.
194. However, the conversation use case advanced by Voxer at trial, in which two users use the broadcast feature to exchange video streams, while it is clunky, is a possible way of using the Facebook technology. And in it there is the capacity to time-shift the media of one message with respect to the other. So I reject this part of Facebook’s case.
195. Another way of looking at this point is to ask whether the claim at least implicitly requires there to be some technical recognition of a semantic connection between the incoming and outgoing messages. That is clearly contemplated by the patent as a possibility (see e.g. the glossary definition of message) but I do not see a justification for reading the attribute into the claim as an essential feature of what a “message” has to be.
196. Before leaving this topic, I will add the following reflection. This aspect of Voxer’s case on what a message is, what an asynchronous message is, and how the claim relates to a conversation, has the result that those concepts have a very wide meaning. They read onto a technique which simply gives multiple users in a network the ability to broadcast live video streams to everyone else. No skilled person reading the patent would think that that was “voxing” or was what the patentee had contributed. Nor would providing VCR functionality to the receivers of such video streams be inventive.

stored at each hop

197. Although the facts can be made to look complicated, they are simple enough. There are servers in the claimant’s relevant application networks which can and do store messages (i.e. the video streams) in a non-volatile manner. One place is what is called Everstore. There are also servers in the relevant application networks which do not store the relevant messages persistently. An example is the PoP servers. The PoP servers will temporarily cache message data as part of the onward transmission process, but that kind of storing is not what the patent is referring to. Facebook also uses a CDN to deliver content to the users. In the CDN, every server has the capacity to store content data but not every server will in fact store particular content.

198. Voxer contend that the right way to look at the Facebook networks is that they consist of a distributed server system in which at least one copy of the message is stored (persistently) somewhere. Therefore, the claim is satisfied. I do not accept that. I find that there are relevant servers in the application networks in issue which had the capability to but do not store the message. They are the servers in Everstore and in the CDN. Therefore, there are hops at which the messages are not stored and the claim's requirement for storage "at each hop" is not satisfied. The PoP servers are not relevant servers.
199. In reaching this conclusion I do not accept Facebook's point that the fact that what is stored in Everstore is the "was live" video rather than the video stream being transmitted live makes any material difference.

Infringement under s60(2)

200. On the conclusions I have reached this point does not arise. It was there to cover a case in which the Facebook systems did not infringe because the ultimate duration for which media was stored in the *tmp* store on the user's device was not under the control of the application software but rather was under the control of the operating system. In that case Voxer contended that the application software would be means essential under s60(2) and the supply of that software would infringe (Facebook denied the point itself but did admit the relevant knowledge). If I had reached the relevant conclusion which meant the point was in issue, I would have found in Voxer's favour on this point.

Equivalents

201. One of the equivalents arguments was advanced on the premise that the requirement for persistently storing the incoming message was not satisfied because the retrieval function in the live broadcast feature was supplemented by network storage. It does not arise. But in case this matter goes further I will say that I was not convinced that the live broadcast feature, if it does not satisfy the claim in that respect, would infringe by the doctrine of equivalents. I will assume in Voxer's favour the second Actavis question. The problem is the first and the third question, looked at together. The skilled reader would understand that in order to allow time-shifting, some persistent storage of the message is necessary since otherwise, after it had been played it would not be available to be time-shifted. There are only two places in which it could be stored – the local device or the network. They are both clearly described in the patent. The reader would understand the similarities and differences between them. If one focusses on the similarities then one might say that it makes no material difference where the data is stored (Q1 in Voxer's favour) but that indicates that it was all the more significant that the patentee, who plainly knew that too, positively chose to claim only local and not network storage (Q3 against Voxer). Or conversely one might focus on the differences and conclude that it does make a material difference to store on the network rather than locally, since (for example) the network may not be available or there may network delays inherent in trying to get the data from the network. Thus either way the equivalents analysis is against Voxer and there would be no infringement. Put another way, and more briefly, this is an example of the "disclosed and not claimed" objection to equivalents.
202. The other equivalents argument advanced by Voxer relates to storage at each hop. Voxer contends that for the full scope of the claim to work, for each message there

needs to be storage in the sending device, storage in the network, and storage in the receiving device. Voxer submitted that it follows that the invention could be implemented using only a single server in the communication network and that Dr Kindberg agreed with that. So he did and I accept that evidence. Voxer then submits that the way Facebook's distributed storage network operates means that whenever the data is required, it is accessible at each server (hop) in the Facebook and Instagram Network from the store which supports that server. I agree.

203. I would answer the first Actavis question in Voxer's favour. The way the Facebook and Instagram Networks are set up in this connection is nothing more than an implementation approach which necessarily arises when one wants to deploy a communication method on a very large scale. It is no doubt a practical necessity for Facebook to use multiple servers geographically distributed around the world. In terms of the way the invention works what is necessary in such a system is that each relevant server ("hop") that has a need to access the messages, is able to do so. That could be implemented in the simple manner described in the patent (storage of identical copies at each hop) but it could also be done by using a distributed storage system which can be accessed by all the relevant servers. The Instagram and Facebook networks achieve substantially the same result in substantially the same way as the invention.
204. I answer Actavis question 2 in Voxer's favour. The functional equivalence would be obvious to the skilled person.
205. Turning to Actavis question 3, Facebook argued that one reason why it must be answered in its favour and against Voxer was because in effect the equivalents argument amounted to concluding that "at each hop" had the same scope as "at least one hop" but that language was added matter. This was said to be supported by the point made by Arnold J as he then was in Akebia v Fibrogen [2020] EWHC 866 at paragraphs 452-545 that the reader must be deemed to be aware of the reasons why a narrowing amendment had been made and that the amendment in effect disclaimed all other ways of achieving the same effect. I agree that the skilled reader will be deemed to know why an amendment has been made. And I can also see that if the reason for the amendment (say) was to remove a particular thing from the scope of the claim because it was prior art, then one can see immediately why that might have a bearing on Actavis Q3 if the patentee was seeking to say that the introduced wording could not be strictly complied with in order to cover the very thing it was introduced to exclude from the claim. However, added matter is a different kind of objection, as I have already explained above. Although it affects the scope of the claim it is in fact concerned with disclosure, which is a different thing. Its purpose as a legal principle is to protect third parties by holding the patentee to their disclosure. Changes in claim scope are dealt with directly by different provisions.
206. In this case the reason for replacing the proposed wording "at least one hop" with "at each hop" was not because "at least one hop" was prior art. It was because "at least one hop" risked being held as a disclosure of new matter, whereas "at each hop" did not. The skilled person deemed to know the reason for the amendment would see that the patentee did regard "at least one hop" as something within the scope of his invention but for technical disclosure reasons was not entitled to write the claim that way. If this indicates anything useful about claim scope (which I believe it does not for the reason already explained), it could indicate to the skilled person that strict compliance with the language "at each hop" was not intended. As I say I believe the right conclusion is that

it is nothing to do with scope, as opposed to disclosure, and so is simply neutral. Another reason for preferring neutrality is to discourage endless failed amendment applications to set up an argument like this one.

207. In my judgment the third Actavis question should be answered in Voxer's favour. The skilled reader would see no reason to think the patentee had intended there to be strict compliance with "at each" hop.
208. Another point made by Facebook was that what the patent discloses the purpose of network storage to be is archiving rather than time-shifting. I agree that that is a purpose of network storage, but I do not agree it would be understood to be limited in that way, nor do I agree that there is a simple distinction between time-shifting and archiving.

Formstein

209. Facebook contended that a so-called **Formstein** defence exists in our law. This is an extension of the **Gillette** defence to a case of equivalents so that if the **Formstein** defence is made out, it leads to the conclusion that there is no scope for equivalents in the given circumstances and the patent's scope must be held to its normal construction.
210. The **Gillette** principle is very well established. It can be stated as being that the patentee cannot validly claim something which was not new or was obvious at the priority date. So in a case in which the alleged infringement can be shown to be obvious over the prior art at the priority date, either the claims must not cover it as a matter of construction or they are invalid. Either way the alleged infringer must win and the patentee must lose. One might think it is obvious that this should apply to equivalents cases too, and so it should. The question is how. While **Gillette** is a useful principle, it has always been recognised in this jurisdiction that, aside from the special case of **Arrow** declarations, it should not be applied directly because everyone – both sides and the public – usually needs to know whether the answer is valid but not infringed or invalid.
211. So imagine a case in which a claim on its normal construction is valid and not infringed, but a defendant's device is (i) found to infringe by the doctrine of equivalents but also (ii) found to be obvious over the prior art. Is the right answer that the claim is infringed but invalid because its proper scope, taking into account equivalents, encompasses something obvious over the prior art; or is it valid but not infringed on the footing that part of the law of equivalents mandates that if these are the facts the equivalents doctrine does not expand the claim? Either answer can be justified logically. Indeed, if the matter was free from authority, given the way the scope of the claim is defined in the EPC itself, one might think the invalidity approach is a purer application of the letter of the law. After all it is how equivalents worked when they were taken into account as part of purposive construction before **Actavis**.
212. **Formstein** is a German case in which the **Gillette** principle was applied in the context of an infringement case. The conclusion was that in such a case the claim was to be held to its normal construction rather than being invalid. However, it is hard to ignore the fact that this made enormous practical sense in the context in which it arose because of the bifurcated nature of the German system. An infringement court would have no jurisdiction to invalidate the claim.

213. ***Formstein*** has been followed in the Netherlands Court of Appeal in ***Eli Lilly v Fresenius*** (Case No C/09/541424). Notably that is not a bifurcated jurisdiction. At para 4.11 the Dutch court in effect treated ***Formstein*** as a fourth question after the three equivalents questions which are more or less the same in every EPC jurisdiction.
214. The United States approach is doctrinally different but comes to what I believe is the same result as ***Formstein*** (see the CAFC's judgments in ***Jang v Boston Scientific*** Nos 2016-1275, 2016-1575 29th Sept 2017 and ***We Care v Ultra-Mark*** 930 F. 2d 1567, 1564-65).
215. So far the UK courts have recognised ***Formstein*** is a possible way forward (see ***Technetix v Teleste*** [2019] EWHC 126 and ***E Mishan v Hozelock*** [2019] EWHC 991 (Pat)) but no UK court has actually had to confront the issue.
216. As things have turned out in this case, I do not have to do so either. However, if I did have to decide the matter, I would hold that the right approach is the ***Formstein*** approach so that the conclusion if the equivalent device lacks novelty or is obvious is that the claim scope must be confined to its normal construction in that respect. I would do so for two reasons. If the claim on its normal construction is valid, then it seems harsh to invalidate it on this ground. What else could the patentee do but write their claim in a way which, normally construed, did not cover the prior art. So that approach promotes certainty. Secondly, since it is clear that other EPC countries work that way, this is a reason in itself for this EPC state to take the same approach.
217. Finally, on the facts, I will say this much. CDNs were common general knowledge. I can see the force of the following argument. The premise would be that the conclusion was reached that either of the prior art based validity attacks fell short only because the system which was obvious involved a CDN and therefore did not "store at each hop" thereby not invalidating the claim on a normal construction. In that case I would hold the CDN equivalents argument was not open to Voxer. However, this does not arise in the present case.

Novelty/Obviousness

218. As things have turned out there is no relevant lack of novelty objection over either item of prior art.
219. The ***Pozzoli*** approach to obviousness is well understood and does not need to be set out. However, to apply it in a straightforward manner requires the parties to set out their case in a helpful way, which neither party did.
220. Before turning to the prior art, one should identify the skilled person or team and the common general knowledge. That has been done. In terms of identifying the inventive concept, this is a case in which it is unhelpful to paraphrase an abstraction over and above the claim language. I have construed the claim above.

Munje

221. *Munje* is entitled "Methods and apparatus for automatically recording Push-To-Talk (PTT) Voice communications for replay". I have addressed the common general

knowledge in the United Kingdom relating to PTT, PoC and what I have called “PTT over cellular” above.

222. As the title suggests, Munje proposes a system for recording the PTT voice communications. The problem Munje starts from is explained in paragraph [0005] as being that PTT communications are generally immediate and unannounced. An end user of the mobile station may be busy or caught “off-guard” and not listening to the initial communication. Therefore, the end user may not hear at least the initial PTT voice communication and so the “talk groups” (i.e. the people to whom the message is addressed) may have to respond to indicate that they did not hear the initial communication. This is inconvenient and wasteful of bandwidth resources. Hence recording the initial message(s) will address that issue. The recording of the message is on the receiving user’s local device. The purpose of that recording is to be replayed at a later date.
223. However, it is also clear that Munje expressly contemplates the idea of recording more than the initial message or two. It is true that the circular buffer memory idea (fig 5) as exemplified appears to hold perhaps two messages. However, the “scrollwheel” or thumbwheel (fig 9) which Munje describes as a way of allowing the user to flip through the recorded messages and select the ones they want to reply to is an indication that more than one or two messages being stored is expressly contemplated by the text itself. More importantly, the skilled team reading Munje would see at para [0074] a proposal to record both the outgoing and incoming PTT messages. The stated purpose of doing so is because this provides in the memory “a more complete history of PTT voice communications”. The outgoing message is stored simultaneously with transmission.
224. In my judgment the skilled team given Munje would see a general proposal to record as much of the incoming and outgoing messages as one might wish to, limited only by the storage available on the local device.
225. The fact Munje is a relatively long document does not detract from this. I reject Voxer’s case that Munje can be summarised as disclosing just the idea of recording on a handset (for later playback) the first few received messages in a PTT conversation. It is more than that, as I have explained.

Identify differences over Munje

226. Neither party approached the case over Munje in a manner which explained with clarity what integers were in issue and what were admitted. Facebook seem to have assumed, without any explanation, that certain aspects were not in issue, but I am not certain I can rely on that. Voxer made some mealy-mouthed admissions about the disclosure of Munje but understanding what is in issue from those is a waste of effort. Rather than trying to grapple with what may have been common ground, it is simpler to work from first principles.
227. Plainly a PTT over cellular system, such as that disclosed in Munje, is a media communication method (feature 1a). Also, plainly what is disclosed by Munje supports a live communication mode (feature 1b) and also a time shifted communication mode (feature 1c). The messages can be exchanged in what amounts to a live conversation or individual messages can be replayed later. Munje’s disclosure relates only to voice.

- There is no reference to video and so feature 1(d) is not satisfied. Features 1(e) and 1(f) are plainly disclosed.
228. Feature 1(g) cannot be dealt with in one go. The first point is “message”. I find that Munje uses messages as required by the claim. The messages are defined by the user.
229. I will address the need for the storing to happen “persistently” separately. Putting that to one side, Munje clearly teaches storing the outgoing message from a communication device in a PTT over cellular system. I find that Munje discloses “progressively encoding, progressively [...] storing on the first communication device and progressively transmitting media of an outgoing message originated on the first communication device over the communication network, as the media is created”.
230. Likewise feature 1(h) cannot be dealt with in one go. Again, I will address the need for the storing to happen “persistently” separately. I will also separate out the question of rendering.
231. Munje teaches locally storing the incoming message to a communication device in a PTT over cellular system. I find that Munje discloses “progressively receiving, progressively [...] storing on the first communication device [...] an incoming message received over the communication network at the first communication device as the media is progressively received [...]”.
232. Munje also discloses the simultaneous storing and rendering (playing) of the incoming message. That was not (I think) in dispute but is in any event explained in Dr Kindberg’s first report (clear from paragraphs 255-256 (option A+D of switches 414 and 416 of Munje Fig 4 (see also [0054])).
233. I turn to “persistently” in both 1(g) and 1(h).
234. The circular buffer mechanism leads to a point on “persistently storing” in claim 1. I reject the idea that as a matter of disclosure Munje is limited to a disclosure to store the messages in RAM. Munje is indifferent to the memory medium itself and circular buffers can be implemented in non-volatile memory. However, this does not mean, as a matter of disclosure, that Munje teaches storage in non-volatile memory, just as “fixing means” does not disclose a nail.
235. Does the fact that either message is stored to be replayed mean it is stored persistently? I would say not. One way of reading Munje is that it teaches storage in RAM because the replay contemplated is in a timeframe associated with the initial message. Moreover, the circular buffer is something which, without more, would automatically wipe earlier messages when a new one came in. I would hold that these two factors mean it is not within what the Voxer patent means by persistently. The use of a circular buffer means that a stored message is liable to be lost to the system without warning and using RAM means that what had been stored would be lost when the power was removed. That applies to feature 1(g) and 1(h).
236. Turning to feature 1(i) (asynchronous messages etc.), Facebook’s closing (para 452) seems to assume this was not in dispute whereas Voxer’s opening (para 215(c)) seems to be on the basis that it is. I find this feature is satisfied by Munje. The messages in Munje are stored separately. They can be listened to in a different time from the time

they were received and therefore an incoming message can be time shifted with respect to an outgoing message by listening to it at a different later time. They are therefore asynchronous messages as required.

237. Feature 1(j) (send without first establishing a connection) was in dispute. I accept Dr Kindberg's evidence that the way the PoC protocol works is by using a client-server model as follows. Consider a client device A in which the user wants to send a message to two recipients B1 and B2. The client device first establishes a connection with a "participating" server SA and passes the message to it. That server SA then passes the message on to a controlling server X, which would then pass two copies of the message forward to a further relevant participating server(s) SB (one for recipient B1 and one for recipient B2). Servers SB pass on the messages to the clients, if the clients are online. This shows that the message is transmitted from client A without needing to first establish a connection with client devices B1, B2 etc. I accept all this. The skilled team given Munje would either know how this all worked anyway or would find out as part of following up Munje.
238. Mr Unitt made the point that the servers could all be the same server, and I accept that. However, it does not alter the conclusion. Implemented as I have described, whether with one server or many, feature 1(j) would be satisfied.
239. Feature 1(k) is not disclosed because Munje does not teach storing the messages on the application server in the network.
240. In relation to Claim 5 (seamless transition) Mr Unitt accepted that it must be implicit in Munje's teaching to retrieve stored messages that the user can exit that interface and return to participate in the live PTT session. I agree. Thus, I would hold claim 5 is disclosed.
241. Claim 10 (queuing when network connectivity is unavailable) is not mentioned in Munje.

Obviousness over Munje

242. I remind myself this is approached without hindsight.
243. The question relates to a United Kingdom based skilled team. I hold that it would not involve an inventive step for such a team, having read Munje, to decide to embark on building a PTT over cellular system implementing the ideas in Munje. Although PTT over cellular was not popular in the United Kingdom, it would be an obvious thing to produce e.g. for the United States market. There is nothing inventive about doing that.
244. I accept Dr Kindberg's evidence that the team implementing Munje would not embark on the project without becoming familiar with the 2004 OMA protocols as well as the future roadmap for that protocol. As at mid-2007 one aspect of the proposals for the future is the idea of including video as well as voice. The team would regard that as an interesting idea worth following up. No hindsight is involved in that. This would be true whether the team decided to approach the project as one based on the PoC OMA protocol itself or as a generic PTT over cellular system. Either is obvious and the team would be well able to put it into practice. Mr Unitt thought video was not obvious and

that the team would not be able to do it even if they thought of it, but I was not persuaded.

245. I accept Dr Kindberg's view that the team would not face any technical difficulties implementing Munje for video messages. At least higher end phones of that era had video cameras, sufficient non-volatile memory to store some video and in many cases the ability for the user to expand the available storage using memory cards. I also note that there is no problem to be solved (to which the Voxer patent is a solution) which would make doing voice and video together inventive. Instead of the individual messages consisting of voice only, they would consist of video, including voice.
246. The fact that Munje's solution might be seen as carefully constructed to be implementable on the existing PoC OMA standards as a self-standing feature makes no difference. A team which looked at Munje's disclosure that way would not feel constrained to follow that approach as they built their own system.
247. The system that the skilled team would set about building would be based on the client - server model. In other words, the system would consist of application software to run on a mobile device such as a phone and server software running on a dedicated network server. The mobiles would communicate with the server via wireless telecommunications networks and the internet.
248. Irrespective of whether the team decided to follow PoC OMA in every respect or not to do so, they would have every reason to use the PoC server architecture and the approach described above which satisfies feature 1j. It is obviously convenient.
249. The system would cater for both "one-to-one" and "one-to-many" group PTT communications. Both are disclosed in Munje (para [0045]) and both were part of the OMA standard at the time, as Dr Kindberg explained. Indeed, Mr Unitt thought the group session idea was the primary use case for Munje.
250. The system would store messages locally on the device. That is what Munje describes. It would apply to both incoming and outgoing messages and would occur simultaneously with their reception at or transmission from the device.
251. There is nothing inventive about storing those messages on that device persistently within the meaning of claim 1. That is because it would be the natural thing for the skilled team to build the system in such a way that the messages would be stored in non-volatile memory and kept for replay for as long as a user wanted to do so. That is particularly so given the suggestion in Munje of storing a more complete history of the communications. There would be no guarantee the history had been stored until it was wanted unless it was stored persistently.
252. Mr Unitt's view was that the team would want to minimise the number of servers and could put the participating and controlling server functions on a single server. I will work on that basis and refer to that as the application server. The system would also persistently store the messages on the application server. There is nothing inventive about that either. Again what matters is that Munje discloses the idea of storing so as to have a more complete history of the session. In that light storage on the network would be a natural thing to do for the following different reasons. First, the mobile devices inevitably do not have unlimited memory and if messages are worth storing for

an appreciable amount of time, storage on the network would be an obvious expedient. By comparison with the mobile devices, the storage available on the network server is for practical purposes unlimited. Second, storage on the network is also obvious when implementing a one to many system, as it would be obvious to allow others to have access to the history which was stored. This would include access by users to whom the messages were directed whose devices were not contactable when the original message was sent.

253. This set up would satisfy the requirement of storage at each hop (since, as would be an obvious thing to do) there would be only one hop.
254. Accordingly Claim 1 is obvious over Munje and so would be claim 5.
255. In terms of claim 10, I am not persuaded there is anything inventive in the idea of queuing messages for progressive transmission when the network is unavailable.
256. Before leaving Munje, I make the following further observations. None of what I have described involves a new way of thinking about communication. To the extent anything is asynchronous or in real time, or involves dividing a session or conversation into pieces, it is simply a function of a conventional PTT system. In terms of motivation, the team would be motivated to build what they would regard as a modern (for 2007) PTT system. It is a technically obvious thing to do, including with video. The fact that in practice video systems of this kind were not introduced until much later does not demonstrate non-obviousness of the concept. Also, I have generally preferred Dr Kindberg's evidence on this topic to Mr Unitt's. One reason why is that I thought Mr Unitt took too narrow a view of what was disclosed by Munje and that weakened the cogency of his opinions.
257. I also look at this another way. Despite the fact that the analysis above involves a number of steps, it is notable that in no case is there a problem to be solved, to which the Voxer patent is a solution, which would make taking that step non-obvious.

Atarius

258. Atarius is a patent application from the well-known telecoms company Qualcomm. It describes the idea of archiving session data in a wireless communications network. The user may want to have a record of the session for retrieval and use later. Session data may consist of anything exchanged via the network during a session, including messages (such as text messages sent in an instant messaging system), voice, video and broadcasts. Data may be continuous (as for voice or video calls), periodic (as for text messages) or sporadic (packet data). The sessions may be many to many, involving multiple users ([0037]).
259. Atarius would be understood as not suggesting that it applies only to certain communication applications. Nevertheless, three particular kinds of communication application are mentioned in [0024]: Instant Messaging (IM), PTT and POC.
260. Atarius is clear ([0044]) that the archiving can be at a terminal (i.e. the local device) and/or on the network. Convenience is a factor in favour of local storage but memory may be limited and so Atarius proposes that the user may initially store data locally at the terminal and then may upload it to the network for archival e.g. if memory was

insufficient or so as to share it with other users. In my judgment the skilled person would understand the idea of local storage here to encompass both the sending and the receiving terminal. Furthermore, it would be understood that the retrieval of the data could be from local storage or from the network, wherever was convenient.

261. In Voxer's infringement case an entire video stream started and stopped by the user consists of a single message. Given that, the fact that there is no reference in Atarius to dividing a conversation into individual messages does not make a difference.

Identify differences

262. Certain aspects can be got out of the way relatively easily. The Atarius proposal includes storage which is persistent, since it is intended for later retrieval, and progressive, since it is stored in real time during the session ([0045]).
263. Feature 1k (at each hop) is satisfied because the skilled person would inevitably, at least, implement Atarius with network storage at a single server.
264. Atarius would be understood by the skilled reader as contemplating a mode of communication which was a live mode of having a conversation – such as a mobile telephone call or PTT conversation.
265. To the extent the data stored amounts to a message, those messages will be asynchronous because they are capable of being reviewed (i.e. for voice to be listened to) at a later time than they were originally heard. I also reject one of Voxer's points, supported by Mr Unitt, that listening to a message which had been archived in this way was not a communication method at all. It plainly is. This retrieval from an archive referred to also means that Atarius is disclosing a time-shifted mode.
266. However, a number of features are not disclosed in Atarius or at least not disclosed together. For example, Atarius does refer to video data and so, when implemented for video it would be a system for communicating voice and video and so satisfy feature 1d. Atarius also refers to instant messaging but it does not follow that Atarius itself discloses the idea of doing instant messaging using video. Nor does it follow that Atarius discloses the combination of PTT or POC with video. These combinations may or may not be obvious but that is a different issue.
267. I am well aware that *Pozzoli* provides a useful framework for considering obviousness but in the end I gave up trying to set out a list of differences between Atarius and claim 1. It is better to move on to consider the various distinct obviousness cases advanced by Facebook.

Obviousness

268. In the end I was not persuaded that Facebook's obviousness cases over Atarius added anything of value to its case over Munje. The most that can be said is that the best case over Atarius is duplicative of the one over Munje. With Atarius one could focus on PTT, apply it, with storage, to a one to many application and then taking the same steps as I have addressed over Munje. However at least Munje is about PTT itself and discloses the combination of PTT and local storage. For Atarius PTT is just one of a number of applications mentioned almost in passing and there is no particular reason

for the skilled person to go down that road. Adding that step onto the others which arise over Munje takes that case over the line such that I am not persuaded the claim would be obvious that way. So I would reject that case anyway.

269. Facebook, supported by Dr Kindberg, ran a different argument over Atarius starting from instant messaging. I accept Dr Kindberg's evidence that incorporating video into instant messaging was well within the technical capabilities of the skilled team. I was not persuaded by Mr Unitt's evidence to the contrary. This applies both to video clips and video conferencing (see below). Instant messaging would satisfy feature 1j.
270. However, I was not satisfied that this thinking leads to the claim without hindsight. I find that the step of following up instant messaging over Atarius is obvious. It would also be obvious to incorporate video. However, without hindsight, the way the skilled team would do that would be by attaching and exchanging pre-recorded video clips. This would not involve progressively transmitting the media of the outgoing message. I have accepted that there were some instant messaging applications, part of the common general knowledge, which supported video conferencing too, but the video conferencing was a distinct part of the application, separate from instant messaging itself. To think of including that as well would be yet another step which would be required and I am not persuaded that further step would be obvious over Atarius. There is nothing in Atarius to lead a skilled team which was following up the reference to instant messaging, to incorporate that sort of video conferencing. They would do nothing more than video clips.
271. Mind you, I do not agree with Voxer that if the team did take that step, they would end up simply with "synchronous" messages and "little more than a simple archive of the entire conference call" (Voxer closing 96(b)). As far as I can tell, if the skilled team did incorporate video conferencing into an instant messaging system using the Atarius archive approach it would exhibit many of the features relied on by Voxer as a basis for the infringement claim. That is because a message can be a whole video stream. The media would be transmitted progressively (and encoded and stored progressively). The messages would be asynchronous because they can be time shifted. It would be obvious to store on the local devices (both incoming and outgoing) for later retrieval and also to store on the network at the single server necessary. It would have a live mode and a time-shifted mode. As an instant messaging system, it would be set up to transmit even though no connection was made at the receiving device, although I am not sure how that would work with video conferencing was explained. However, it is not necessary to examine this further since I was not persuaded the video conference aspect was obvious.
272. Even if claim 1 was obvious over Atarius, I am not convinced claim 5 would be. It would not be obvious to build VCR functionality into video conferencing.
273. However, if claim 1 was obvious, claim 10 would also be obvious. There is nothing inventive over claim 1 about queuing for progressive transmission if the network is unavailable.

Insufficiency

274. The insufficiency points which fall to be decided are:

- i) An argument that the term “asynchronous messages” is wholly unclear;
 - ii) A point that the skilled person is not able to “demarcate” videos into messages;
 - iii) A general submission that the patent is no more enabling than the prior art, which spawns various alleged squeezes.
275. Facebook’s point on asynchronous messages was a submission of insufficiency by uncertainty, which in effect has already been rejected. The skilled person would understand what the patent is getting at, including the use of asynchronous messages to give rise to a live mode and time-shifted mode. I reject the idea that the skilled person would have any difficulty putting the invention into practice because of any doubt about what “asynchronous messages” are. To the extent Dr Kindberg gave admissible evidence to the contrary, I do not accept it.
276. The issue about demarcating messages in video has more substance. The starting point is to consider a conversation between two people by voice alone over a telecommunications system. To the individuals it may appear as a continuous exchange, however the evidence of the experts was clear that the skilled team would know, as a matter of common general knowledge, that there were various ways of dividing up the audio into discrete messages from the point of view of the telecommunications system. For example, in PTT the messages can be defined by a user’s command, e.g. holding down a button as the user speaks. Another approach, not limited to PTT, would be for the system to detect silences greater than a prescribed threshold, for example between spoken words or phrases or by another participant starting to speak.
277. However demarcating video into discrete messages is a different problem. After all video is effectively continuous. In cross-examination Mr Unitt agreed with Dr Kindberg that apart from a technique like PTT with a start and stop button which applied to the video to demarcate the message boundaries, it is difficult to see how the system could automatically demarcate messages within a live stream of video. There will be specific examples in which the video stream has a natural break – such as half-time in a football match. However in the general case, the only way the skilled team would be able to put the invention into practice for a video stream is by providing a user interface (such as a button) so that the user marks the beginning and end of the “message”.
278. To the extent this case was pleaded at all, Facebook’s point was advanced as a squeeze over the prior art [Grounds of Invalidity 6(d)]. In other words it would not be open to Voxer to rely on any difficulties in demarcating video streams aside from a user’s command, when taking obvious steps over Munje or Atarius. I have approached the case that way and in that sense the squeeze has done its job. This point also explains why Voxer’s infringement case treats the entire video streams as single messages.
279. However, in closing Facebook also sought to expand the scope of this plea by submitting that this point means that claim 1 is not enabled across its breadth and so is invalid for insufficiency in any event, citing *Regeneron v Kymab* [2020] UKSC 27. That point was not pleaded and I reject it on that ground. I believe it is not well-founded anyway for the following reason. The skilled team is able to demarcate any video stream by the user command technique. Nevertheless, it is true that there will be other

ways of demarcating video streams into “asynchronous messages” which the patent does not enable. As I sought to explain in *Illumina v Latvia MGI* [2021] EWHC 57 (Pat) paragraphs 249 to 279, the requirement that a range has to be enabled across its whole scope applies to ranges relevant in the *Regeneron* sense and not to other ranges. In my judgment, the term “asynchronous messages” in claim 1, to the extent it is a range at all, is not a range relevant in the *Regeneron* sense and so even if this point was open to Facebook, I would reject it.

280. Turning to the final category of insufficiency relied on, the main point was that at various points in the argument Voxer, supported by Mr Unitt, had sought to emphasise difficulties in implementing video, such as in a PTT system starting from Munje. I was not persuaded there was any relevant technical difficulty for the skilled team implementing video telecommunications systems whether over the prior art or in putting the patent into practice. That applies whether the team is taking forward PTT, broadcasting, instant messaging or video and voice conversations. There might have been some problem with persuading standard setters to modify standards but (a) I was not convinced and (b) that kind of difficulty is not relevant in this case.
281. On claim 5 Facebook took a point that it is not clear how one can transition from a time-shifted viewing mode to a live form where the live transmission has already come to an end. No doubt that is true but no skilled person would think that was a requirement of the claim. The ability to transition to live mode depends on the live mode still being in progress to transition to.
282. I think Facebook may have been seeking to make a different point based on one of the issues in the infringement debate that the “was live” stream of a broadcast was different in some ways from the live stream. So I think it may be being said that even if, when the live broadcast is still going, a user watching the “was live” stream can fast forward up to the live head of the live stream, that is not within the claim because the two streams are different. I have rejected that already, but whether that rejection is right or wrong there is no insufficiency revealed.

Conclusion

283. I conclude that Voxer’s patent EP (UK) No. 2 393 259 is not infringed by any of Facebook’s live broadcast features implemented in the iOS Facebook App, the iOS Instagram App or the Facebook website on a personal computer.
284. The amendments proposed to the patent in the re-amended claim set of 25th February 2021 are allowed. The claims are novel, are not obvious over WO 2006/121550 (Atarius), nor are they insufficient. However as amended the patent is invalid for obviousness over US 2006/0003740 A1 (Munje).

Annex A Re-Amended Claims (25th Feb 2021)

The red amendments were in the claim set of 20th Nov 2020. The green amendments were proposed on 25th Feb 2021. The green struck through amendments were then dropped by agreement [*Facebook closing para 26(b) fn 13*].

Claim 1

A media communication method, which supports a live communication mode and at least one time-shifted communication mode, for communicating both voice and video media on a first communication device (13) over a communication network (14), comprising:

progressively encoding, progressively and persistently storing on the first communication device (13) and progressively transmitting media of an outgoing message originated on the first communication device over the communication network, as the media is created; and

progressively receiving, progressively and persistently storing on the first communication device (13) and progressively rendering media of an incoming message received over the communication network at the first communication device as the media is progressively received in a real-time rendering mode,

wherein the outgoing message and the incoming message are asynchronous messages (such that the media of an incoming message may be time-shifted with respect to the media of the outgoing message) that are transmitted over the communication network from the first communication device to the second communication device and received over the communication network at the first communication device from the second communication device without first establishing a connection over the communication network between the first communication device and the second communication device

and wherein the outgoing message and the incoming message are stored and transmitted at at least one each hop along a path over the communication network.

Claim 1A

The media communication method of claim 1 which supports a live communication mode:

wherein when network conditions are poor, the quality of the media for transmission is intentionally reduced to the point where it is good enough to be rendered upon receipt;

and wherein an exact copy of the media is eventually delivered over time.

Claim 2

The media communication method of claim 1 ~~or claim 1A~~ which supports a live communication mode wherein:

the outgoing message and the incoming message are near synchronous at the first communication device; and

the first communication device utilises the real-time rendering mode to progressively render media of the incoming message.

Claim 3

The media communication method of any of the preceding claims which supports a time-shifted communication mode wherein:

the outgoing message and the incoming message are near synchronous at the first communication device; and

the first communication device progressively renders media of an incoming message out of storage on the first communication device sometime after the media was received from the second communication device.

Claim 4

The media communication method of claim 1, ~~1A~~ or 2 which supports a time-shifted communication mode wherein:

the outgoing message and the incoming message are time-shifted by storage in the communication network so as to not be near synchronous at the first communication device; and

the first communication device utilises the real-time rendering mode or the time- shifted rendering mode to progressively render media of the incoming message.

Claim 5

The media communication method of any preceding claim wherein the first communication device can seamlessly transition back and forth between the live communication mode and the time-shifted communication mode.

Claim ~~2~~ 6

The media communication method as claimed in any preceding claim ~~4~~, wherein the media of the outgoing message originated on the first communication device and the media of the incoming message received over the communication network are stored as time-indexed messages.

Claim ~~3~~ 7

The media communication method as claimed in any of claims ~~1-5~~, wherein the media originated on the first communication device and the media received over the communication network are segmented into time-indexed messages and the messages are threaded into conversations.

Claim ~~4~~ 8

The media communication method as claimed in claim ~~7~~ 3, wherein each message is assigned

an attribute indicating the conversation it belongs to.

~~Claim 5~~

~~The media communication method as claimed in claim 1, wherein the method supports a live communication mode wherein media originated at the first communication device is progressively transmitted and media received over the communication network is progressively received during the live communication mode.~~

~~Claim 6~~

~~The media communication method as claimed in claim 1, wherein the method supports a time-shifted communication mode in which media received over the communication network is rendered out of storage on the first communication device sometime after the media was received from the second communication device during the time-shifted communication mode.~~

Claim ~~7~~ 9

The media communication method as claimed in any preceding claim, further comprising providing a user interface on the first communication device which enables a user to generate media or review media from storage.

Claim ~~8~~ 10

The media communication method as claimed in any preceding claim, wherein media that is created by the communication device when network connectivity is unavailable is queued for progressive transmission from persistent storage as soon as network connectivity is available.

~~Claim 9~~

~~The media communication method of any preceding claim, wherein the outgoing message is transmitted, and the incoming message is received, by storage and transmission at each hop along a path over the communication network.~~

Claim ~~10~~ 11

The media communication method as claimed in any preceding claim, wherein the method supports a live communication mode at the second communication device wherein the second communication device progressively renders media of the outgoing message as media of the outgoing message is received from the first communication device over the communication network.

Claim ~~11~~ 12

A communication device (13) having a client application (12) stored thereon which, when executed by the device performs the method of any of claims 1 to ~~10~~ 8.

Claim ~~12~~ 13

A medium (146) storing a client application executable by a processor (142) to carry out the method of any of claims 1 to 10 8.